Tuesday

UNIT-II

GoodLuck Page No.
Date 02,04,2024

Assignment No:-2

Q.] Ten Individuals are choosen at random from a population & their height are found to be in inches 63, 64, 63, 65, 66, 69, 69, 70, 70, 71. Discuss the proposal that mean height in universe is 65 inches (take $\alpha = 0.05$).

Ans → • Here, we can apply T-test :-

$\therefore n = 10$

• step 1: Define Hypothesis :-

• Null Hypothesis, $H_0 = (\mu) = 65$ (population of mean)

• Alternative Hypothesis $(H_1): (\mu) \neq 65$.

• step 2: ~~Calcult~~ Compute T-statistics :-

$$\therefore t = \frac{\bar{x} - \mu}{S} \sqrt{n}$$

• calculate S.D of Sample :-

x	$(x-\bar{x})$	$(x-\bar{x})^2$
63	-4	16
63	-4	16
64	-3	9
65	-2	4
66	-1	1
69	2	4
69	2	4
70	3	9
70	3	9
71	4	16
$\Sigma x = 670$		$\Sigma(x-\bar{x})^2 = 88$

GoodLuck Page No.
Date

$\therefore \bar{x} = \frac{\Sigma x}{N}$

$= \frac{67\cancel{0}}{1\cancel{0}}$

$\therefore \bar{x} = 67$

$\therefore$ Variance of Sample (s^2) :-

$\therefore s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1}$

$= \cancel{\frac{88}{101}} \ \frac{88}{10-1}$

$= \frac{88}{9}$

$\therefore s^2 = 9.77$

$\therefore S.D = \sqrt{9.77}$

$\therefore S.D = 3.1257$

$\therefore t = \frac{\bar{x}-\mu}{s} \sqrt{n}$

$= \frac{67-65}{3.1257} \times \sqrt{10}$

$= 0.6399 \times \sqrt{10}$

$\therefore$ Cal $|t|$ $= 2.0235$

$\therefore$ Degree of Freedom $(\nu) = n-1$

$= 10-1$

$= 9$

- Step 3 :- Decision :-

$\therefore \alpha = 0.05$

$\therefore \nu = 9$

$\therefore t(9, 0.05) = 2.262$

$\therefore |t|cal < t(0.05, 9)$

$\therefore 2.0234 < 2.262$

$\therefore$ Ho is accepted.

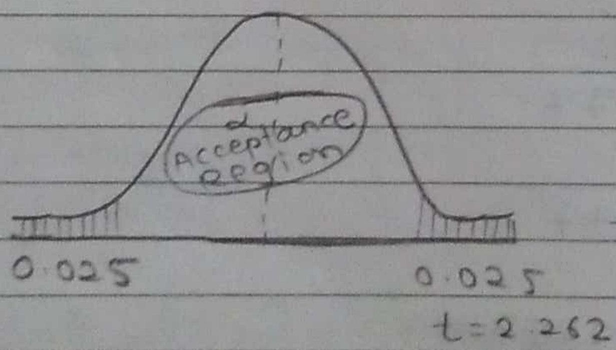


fig. Hypothesis.

Q.] Below are given the gain in weights (in kgs) of dogs fed on two diets A & B

	Gain in Weight
Diet A	25, 32, 30, 34, 24, 14, 32, 24, 30, 31, 35, 25
Diet B	44, 34, 22, 10, 47, 31, 40, 30, 32, 3, 18, 21, 35, 29, 22.

Test if two diets differ significantly as regards their effect on increase in weight (take $\alpha = 0.05$)

1] Define Hypothesis :-

$\therefore H_0 : \mu x_1 = \mu x_2$ (Null Hypothesis)

$\therefore H_1 : \mu x_1 \neq \mu x_2$ (Alternative Hypothesis). (Two tailed Test)

2) Calculate t:-

$$\therefore t = \frac{\bar{x_1} - \bar{x_2}}{S} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

Diet A	$(x_1 - \bar{x_1})$	$(x_1 - x_1)^2$	Diet B	$(x_2 - \bar{x_2})$	$(x_2 - \bar{x_2})^2$
25	-3	9	44	~~91667~~14	196
32	4	16	39	4	16
30	2	4	22	-8	64
34	6	36	10	-20	400
24	-4	16	47	17	289
14	-14	196	31	1	1
32	4	16	40	10	100
24	-4	16	30	0	0
30	2	4	32	2	4
31	3	9	35	5	25
35	7	49	18	-12	144
25	-3	9	21	-9	81
			35	5	25
$\Sigma x_1 =$		$\Sigma(x_1 - x_1)^2 =$	29	-1	1
336		380	22	-8	64
			$\Sigma x_2 =$ 450		$\Sigma(x_2 - \bar{x_2})^2 =$ 1410.

$$\therefore \bar{x_1} = \frac{\Sigma x_1}{N}$$

$$= \frac{336}{12}$$

$$\therefore \bar{x_1} = 28$$

$\therefore \cancel{x^2}$

$$\therefore \bar{x_2} = \frac{\Sigma x_2}{N}$$

$$= \frac{450}{12}$$

$\therefore \cancel{\bar{x_2} = 34.8333}$

$$\therefore \bar{x_2} = 30$$

GoodLuck Page No.
Date

$$\therefore S = \sqrt{\frac{\Sigma(x_1-\bar{x}_1)^2 + (x_2-\bar{x}_2)^2}{n_1+n_2-2}}$$

$$= \sqrt{\frac{380+1410}{12+15-2}}$$

$$= \sqrt{\frac{1790}{25}}$$

$$\therefore S = 8.4617$$

• Calculate Value of t :-

$$\therefore \text{Cal}|t| = \frac{\bar{x}_1 - \bar{x}_2}{S} \times \sqrt{\frac{n_1 n_2}{n_1+n_2}}$$

$$= \frac{28-30}{8.46} \times \sqrt{\frac{12\times15}{12+15}}$$

$$= \frac{-2}{8.46} \times \sqrt{\frac{180}{27}}$$

$$\therefore \text{Cal}|t| = 0.6104$$

• Calculate Degree of Freedom,

$$\therefore df(\nu) = n_1+n_2-2$$
$$= 12+15-2$$
$$\therefore df(\nu) = \underline{25}$$

• Step 3: Decision :-

∴ Take α as 0.05

$\therefore |t| < t(0.05, 25)$

$\therefore 0.6104 < 2.06$

$\therefore$ H_0 is accepting

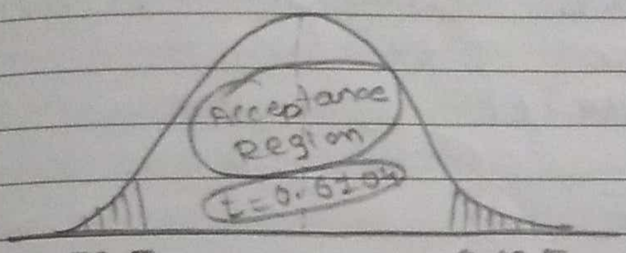


fig. Hypothesis

6) A sports governing body wants to test whether a supplement used by professional athletes to be banned for increasing levels of testosterone in body. The levels of testosterone in picograms / millilitre of ten atheletes were tested before & after taking the supplement, with the results summarised in the following table,

i	1	2	3	4	5	6	7	8	9	10
Before (pg/ml)	65.83	111.15	106.11	91.12	97.4	135.89	69.45	83.33	157.88	74.69
After (pg/ml)	77.92	129.27	109.72	97.68	124.37	147.12	71.16	81.27	164.16	79.51

Perform a hypothesis test using 1% significance level to decide whether or not the supplement should be banned.

Ans → • Step 1: Define Hypothesis :-

∴ H_0 : Supplement should be banned.

∴ $H_0 = \mu x_1 = \mu x_2$

∴ $H_1 = \mu x_2 \neq \mu x_2$ → (Two tailed test).

• Step 2 :- Calculate |t|

$$\therefore |t| = \frac{\bar{d}}{s}\sqrt{n}$$

x_1	x_2	$d = x_1 - x_2$	$(d-\bar{d}) = (d-d-(-2))$ $= d+2$	$(d-\bar{d})^2$
65.83	77.92	-12.0900	-10.0900	101.8081
111.15	129.27	-18.1200	-16.1200	259.8544
106.18	109.72	-3.5400	-1.5400	~~1.5~~ 2.3716
91.12	97.68	-6.5600	-4.5600	20.7936
97.43	124.37	-26.9400	-24.9400	622.0036
135.89	147.12	-11.2300	-9.2300	85.1929
69.45	71.16	-1.7100	0.2900	0.0841
83.33	81.27	2.0600	4.0600	16.4836
157.88	164.16	-6.2800	-4.2800	18.3184
74.69	79.51	-4.8200	-2.8200	7.9524
		$\Sigma d = -89.2300$		$\Sigma(d-\bar{d})^2 = 1134.8627$

$$\therefore \bar{d} = \frac{\Sigma d}{n} = \frac{-89.2300}{10} = \underline{-8.9230}$$

$$\therefore S = \sqrt{\frac{\Sigma(d-\bar{d})^2}{n-1}}$$

$$= \sqrt{\frac{1134.8627}{10-1}}$$

$$= \sqrt{\frac{1134.8627}{9}}$$

$$= \sqrt{126.0959}$$

$$\therefore S = 11.2292$$

• Test statistics:-

$$|t| = \frac{\bar{d}}{S}\sqrt{n}$$

$$= \frac{-8.9230}{11.2292}\sqrt{10}$$

$$= -0.7946 \times \sqrt{10}$$

$$\therefore |t| = 2.5127$$

• Step 3: Decision

$$\therefore d.f(\nu) = n-1$$

$$df(\nu) = 9$$

$$\therefore t(0.11, 9) = 4.781$$

$$\therefore 2.5127 < 4.781$$

$\therefore H_0$ is accepted.

Q.) What is meant by t-test ? Explain four type of T-test?

Ans→ *T-Test :-

• A t-test (also known as student's t-test) is a tool for evaluating the means of one or two populations using hypothesis testing. A t-test may be used to evaluate whether a single group differs from known value (a one-sample t-test), whether two groups differ from each other (an independent two-sample test), or whether there is a significant difference in paired measurement (a paired, or dependent samples t-test).

• A t test is appropriate to use when we have collected a small, random sample from some statistical "population" and want to compare the mean from our sample to another value. The value for comparison could be fixed value (e.g. 10) or the mean of second sample.

• For Ex, if our variable of interest is the average height of sixth graders in our region then we might measure the height of 25 or 5 randomly - selected ~~with~~ sixth graders. A t-test could be used to answer questions such as, "Is the average height greater than four feet?"

• t- Test Assumptions:-

• While t-tests are relatively robust to deviation from assumptions, t-tests do assume that:-

- The data are continuous.
- The sample data have been randomly sampled from a population.
- There is ~~no~~ homogeneity of variance. (variability of data in each group is similar).
- The distribution is approximately normal.
- For two-sample t-test, we must have independent samples. If samples are not independent, then a paired t-test may be appropriate.

• Hypothesis Testing:-

• one & two-tailed tests are ways to identify the relationship between the statistical variables. For checking the relationship between variables in a single direction (Left or right direction), we use a one-tailed test. A two-tailed test is used to check whether the relations between variables are in any directions or not.

*one-Tailed Test:-

• A one-tailed test is based on uni-directional hypothesis where the area of rejection is the on only one side of sampling distribution. It determines whether a particular population parameter is larger or smaller than the pre-defined parameter. It uses one single critical value to test data.

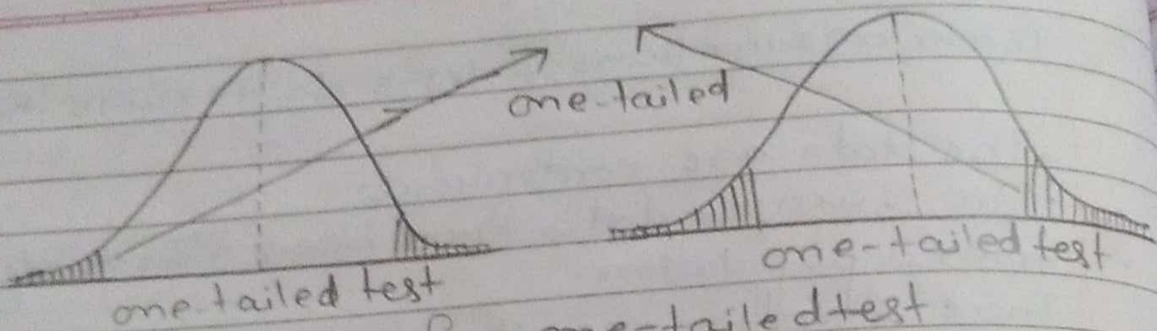


fig- one-tailed test

- Null Hypothesis (H_0): Where represents a parameter (e.g. population mean) & 'θ_0' is a specific value.

- Alternative Hypothesis (H_1):-

- For a right-tailed test: $\theta > \theta_0$.
- For a left-tailed test: $\theta < \theta_0$.

- Test statistic :- Depending on type of test & the distribution, the test statistic is computed (for ex:- t-score for normal distribution)

- Decision Rule :- If the test statistics falls in the critical region, reject the null hypothesis in favor of alternative hypothesis.
- Ex :- Effect of participants of students in coding competition on their fear level.

- H_0: There is no important effect of students in Coding competitions on their fear level.
- The main intention is to check the decrease fear level when student participate in a coding competition.

* Two-tailed test :-

• A two-tailed test is also called a non-directional hypothesis. For checking whether the sample is greater or less than a range of values, we use the two-tailed. It is used for null hypothesis testing.

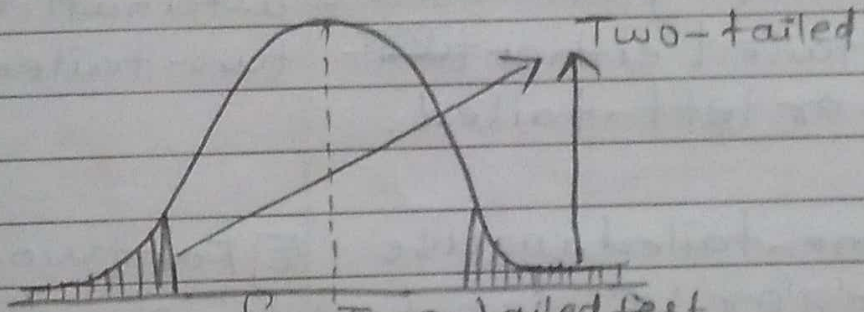


fig Two-tailed test

• Null Hypothesis (H0) :- where represents a parameter (e.g, population mean) and @0 is a specific value.

• Alternative Hypothesis (H1): $\theta \neq \theta_0$

• Test statistic :- Compute the test statistics as appropriate for the distribution (for example t-score for normal distribution).

• Decision Rule :- If the test statistics falls in either tail of the distribution's critical region, reject the null hypothesis in favor of the alternative hypothesis.

• Ex :- Effect of new bill pass on loan of farmers.

• H0: There is no significant effect of new bill passed on loans of farmers. New bill passes can affect in both ways either increase or decrease the loan of farmers.

GoodLuck Page No.
Date

*Difference between one & two-tailed Test

One-Tailed Test	Two-tailed Test
1] A test of any statistical hypothesis, where the alternative hypothesis is one-tailed either right tailed or left-tailed.	1] A test of a statistical hypothesis, where the alternative hypothesis is two-tailed.
2] For one-tailed, we use either '>' or '<' sign for the alternative hypothesis	2] For two-tailed, We use '≠' sign for the alternative hypothesis.
3] When the alternative hypothesis specifies a directions then we use a one-tailed test.	3] If no directions is given then we will use a two-tailed test.
4] Critical region lies entirely on either right side or left-side of the sampling distributions.	4] Critical region is given by the portion of the area lying in both the tails of the probability curve of test statistic.
5] Here, the entire level of Significance (α) i.e 5% has either in the left tail or right tail.	5] It splits the level of Significance (α) into half.
6] It checks relation between variables in single direction.	6] It checks relation between the Variables in any direction

7] Rejection region is either from left side ~~of the sampling~~ or right side of sampling distribution	7] Rejection region is from both sides i.e left & right of the sampling distribution.
8] It is used to check whether the one mean is different from another mean or not.	8] It is used to check whether two mean different from one another or not.

* Types of T-Test :-

- There are three types of t-test :-

1] One Sample T-Test.

2] Independent Sample T-Test.

3] Paired Sample T-Test.

1] One Sample T-Test :-

- As the name implies, this test is used when we have one data set for sample & we need to determine whether this data set belongs to a particular population or not. The mean value for the population data must be known in this case. The formula to determine T-value, in this case, is as follows:-

$$\therefore t = (\bar{x} - \mu) / (\sigma / \sqrt{n}).$$

where,
- t is the t-value.
- $\bar{x}$ is the sample mean.
- μ is the population mean.
- σ is the sample standard deviation &
- n is sample size.

- steps to calculate T-value for one Sample T-Test

- To perform the one sample T-test, the steps listed below are generally followed:

- Step 1:- State a null hypothesis & an alternative hypothesis. The null hypothesis assumes that the sample mean & the known population mean (μ) are equal, while the other assumes that sample mean is different from the population mean.

- Step 2:- Define values for the level of significance (α) & the degree of freedom (df). The degree of freedom equals $(n-1)$ for this case.

- Step 3:- Calculate t-value using formula stated above by putting all the known values of sample mean ($\bar{x}$). Sample standard deviation (σ), the population mean (μ) & the sample size (n).

- Step 4:- Determine the associated t-value using t-distribution table.

- Step 5:- Compare t-value to calculate t-value. If the calculated t-value is greater than tabular value, reject the null hypothesis & conclude that the sample mean is significantly different from the population mean. otherwise, ~~for~~ conclude that there is no significant difference between the sample mean & population mean

2) Independent Sample T-Test:-

- As the name suggests, an independent sample T-test is used when we need to compare the stastical means of two independent samples or groups. It helps us determine whether there is significant difference between the means of the two groups. If there is significant differences, it suggests that the groups likely have different population mean, otherwise, they have the same population means.

- T-Test for Independent ~~values:~~ samples:-

- The steps listed below are generally followed to perform this test.

- step 1:- state a null hypothesis & an alternative ~~hypothesis~~ hypothesis. The null hypothesis assumes that the mean of two groups are equal ($\overline{x_1} = \overline{x_2}$), while the other assumes that means of two groups are significantly different ($\overline{x_1} \neq \overline{x_2}$).

- step 2:- Define the values for the level of significance (α) & the degrees of freedom (df). The degree of freedom ~~(df~~ equals ($n_1 + n_2 - 2$) in this case.

- Step 3:- Calculate the t-value from the formula after obtaining the required data related to each other

$\therefore t = ((\bar{x}_1 - \bar{x}_2)/S) * \sqrt{n_1 n_2 / n_1 + n_2}$

- Step 4:- Find the critical t-value from a t-distribution table with the corresponding degrees of freedom & level of significance.

- Step 5:- If the calculated t-value is greater than the critical t-value, then reject the null hypothesis. This indicates that there is significant difference between the means of two groups. Otherwise, the null hypothesis is not rejected. And, this suggests that there is no significant difference between the means of two groups.

3) Paired Sample T-Test:-

- The paired sample t-test is used when we want to compare the means of two related groups or samples. For ex, we may use this test to compare the average scores of players of an athletic team before & after training program. To calculate the t-value in this case, the following formula is used,

D = Difference between two paired samples.

d_i = The ~~sample~~ ith observation in D.

n = The sample size.

$\bar{d}$ = The sample mean of the differences.

$\hat{\sigma}$ = The sample S.D of differences.

T = The critical value of t-distribution with (n-1) degrees of freedom.

t = The t-statistics (t-statistics test) for sample distribution.

GoodLuck | Page No. | Date

- The four steps are listed below:

1] Calculate the sample mean.
$\bar{d} = d_1 + d_2 + \cdots + d_n$

2] Calculate the sample S.D.
$\therefore S = \sqrt{\Sigma (d - \bar{d})^2 / n - 1}$

3] Calculate t value:-

$t = (\bar{d} / s)\sqrt{n}$.

- Steps for paired Sample T-test:-

- Following are steps to perform this type of T-testing

- Step 1:- State the null hypothesis which assumes that there is no significant difference between the statistical ~~which assumes that there is a~~ means of the paired observations ($\mu d = 0$) while the alternative hypothesis assumes that there is significant difference between the statistical means of the paired observations ($\mu d \neq 0$).

- Step 2:- Match each observation in one group with a corresponding observation in other group.

- Step 3:- Calculate the difference between each paired observation & then calculate the mean of differences & the sample S.D of differences & further more, calculate t-value from the formula.

- Step 4:- Obtain the critical t-value from a t-distribution table corresponding to the choosen level of significance (α) & degree of freedom (df). The degree of freedom (d) equals (n-1) in this case.

• step 5:- If the calculated t-value is greater th the critical t-value, then reject the null hypoth This indicates a significant differences in th Sample before & after the intervention. Otherwi it can be concluded that there is no significan difference in the sample before & after the intervention.

Q.] What is students t-distribution? Explain the conditions to apply students t-Test.

Ans→ *Student's t-distribution :-

• Student's t-distribution, also known as the t-distribution, is a probability distribution that used in statistics for making inferences about the population mean when the sample size is small or when the population S.D is unknown. is similar to standard normal distribution (Z-distribution), but it has heavier tails. Theoratical work on t-distribution was done by W.S. Gusset has published his findings under the pen name "student". That's why it is called a student's t-test. t-score represents the number of S.D the sample mean is away from the population mean

• T-Score :-

The 'T-score' also known as the t-value or t-statistics, is a standardized score that quantif how many standard deviations a data point or a sample mean is from the population mean. It is commonly used in statistical hypothesis testing, particularly in scenario where the sample size is small or the population S.D is unknown.

· The formula for calculating the T-score in the context of a t-distribution is given by:

∴ $t = (\bar{x} - \mu/s) * \sqrt{n}$ where,

∴ t = t-score
∴ $\bar{x}$ = sample mean.
∴ μ = population mean.
∴ S = S.D of sample.
∴ n = Sample size.

· As we know, we use t-distribution when the S.D of the population is unknown & the sample size is small. The formula for t-distribution looks very similar to the normal distribution; the only difference is that instead of S.D of population, we will use the S.D of sample.

· When to use the t-Distribution:-

· Student's t Distribution is used when:-

a) The sample size is 30 or less than 30.

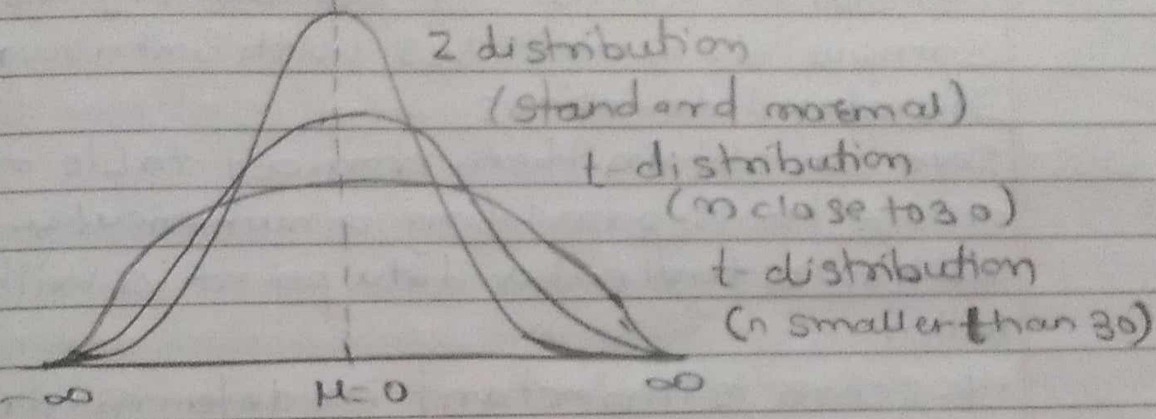
b) The population standard deviation (σ) is unknown.

c) The population distribution must be unimodal & skewed.

*Properties of t-distribution :-

- The variable in t-distribution ranges from $-\infty$ to $+\infty$ ($-\infty < t < +\infty$).
- t-distribution will be symmetric like normal distribution if the power of t is even in the probability density function (pdf).
- For large values of v (i.e. increased sample size) the t-distribution tends to a ~~higher special~~ standard normal distribution. This implies that for different v values, the shape of t-distribution also differs.
- The t-distribution is less peaked than the normal distribution at center and higher peaked in the tail. ~~d~~ observe that From the above diagram, one can ~~d~~ observe that the red & green curves are less peaked at the center but higher peaked at the tail than the blue curve.
- The value of y (peak height) attains highest at $\mu = 0$ as one can ~~des~~ observe the same in above diagram.
- The mean of distribution is equal to 0 for $v > 1$, where v = degree of freedom, otherwise undefined.
- The median and mode of the distribution is equal to 0.
- The variance is equal to $v/v-2$ for $v-2$ & ∞ for $2 < v \leq$ otherwise undefined.
- Degree of freedom ~~refer~~ refer to the number of independent observations in set of data. When ~~so~~ estimating a mean score for a proportion from single sample, the number of independent observations is equal to sample size minus one.
- Hence, the distribution of t-statistic from samples

of size 10 would be described by a t-distribution having 10-1 or 9 degree of freedom. Similarly, a t-distribution having 15 degrees of feedom would be used with a sample size 16.



*t-distribution table:-

• t-distributiontable gives the t-value for different level of significance and different degree of feedom. The calculated t-Value will be ~~ex~~ compared with the tabulated t-value. For ex, if one is performing a student's t-test & for that performance, he has taken a 5% level of significance & he got or calculated & t-value & he has taken his tabulated t-value & if the calculated t-value is ~~hight~~ higher than the tabulated t-value, in that case, it will say that there is a significant difference between the population mean & the sample mean at 5% level of significance & if vice versa then, in that case, it will say that there is no significant difference between the population means & the sample mean at 5% level of significance.

*t-score & p-values:-

• t-score :-

• It represents the deviation of data point from mean in a t-distribution, expressed in terms of standard deviations. Particularly useful for s Sample sizes or cases with unknown populat standard deviations.

• We can obtain them from a t-table or throug online tools, providing a numerical ~~of~~ measure of how a typical a data point is within the di bution.

• t-score is important in determining confide intervals, aiding in estimating the range withi which the true population parameter is likely to fall The critical value of t is ~~e~~ integral ~~is~~ in co dence interval calculations, guiding the determi of upper & lower bounds.

• P-value :-

• The p-value (probability value) is a statistical m sure that helps acces the evidence again a null hypothesis.

• P-value describes the likelihood of data occurring the null hypothesis were true.

• We can use statistical Software to directly obtai the p-value associated with the calculated t-sco or we can use the t-table, which provides c values for different levels of significance and degrees of freedom and the column correspo to our t-score to get the p-value.

1) Testing for the Hypothesis of population Mean
T-distribution are commonly used in hypothesis tests regarding the population mean. This involves accessing whether a sample mean is significantly different from a hypothesized population mean.

2) Testing for the Hypothesis of the difference between two mean:-
T-tests can be employed to examine if there is a significant difference between the means of two Independent Samples. This can be done under the assumption of equal Variance or when variances are unequal. In Scenarios where samples are not independent, such as paired or dependent samples, t-tests can be assess the Significance of mean difference between related observations.

3) Testing for the Hypothesis about the coefficient of correlation:-

T-distribution play a role in hypothesis testing related to correlation coefficients. This includes situations where the population correlation coefficient is assumed to be zero ($\rho=0$) or when testing for a non-zero correlation ($\rho\neq0$)

*Difference Between T distribution and Normal distribution:-

T-Distribution	Normal Distribution.
1] T-Distribution is defined by it's degree of freedom which itself depends upon the sample size.	1] Normal distribution is defined by it's mean & S.D.
2] T-distribution is used when the sample size is small.	2] Normal distribution is used when we have large no data points lie near the mean of distributions.
3] It has heavier tail than normal distribution which means more datapoints are away from mean of the distribution.	3] Normal distribution has lighter tail than T-distribution which means more data points lie near mean of distribution.
4] We use T-distribution in hypothesis testing when the S.D of population is unknown.	4] Normal distribution is used when the S.D is known.
5] T-Distribution has a larger range of critical values as compared to the normal distributions as this distributions has heavier tails.	5] Normal distribution has smaller range as compared to t-distribution.

unday

GoodLuck Page No
Date 07 | 03 | 2024

Q.] Explain degree of Freedom with proper example?

ans→ *Degree of Freedom :-

- The degrees of freedom that are mathematical concepts to statistical calculations ~~pre~~ represent the number of variables that have the freedom to vary in a calculation.
- calculating degrees of freedom can help ensure the validity of t-tests & highly f-tests, among other tests. These tests are often used to compare data that has been detected with data that would be expected if a particular hypothesis were true.
- The fact that the statistical degrees of freedom indicating the number of values in the final calculation is allowed to vary means that they can contribute to the validity of result. Although the number of observations & parameters to be measured th depends on the size of sample, or the number of observations & the parameters to be measured, the degree of freedom in the calculation is usually equal to the value of observation minus the number of parameters
- This means that for larger sample size, there are degrees of freedom ~~ava~~ available.
- For Example :-
- Mention that, we have seven shirts that we can wear for a week, & we decide to wear each shirt only once a week.
- On Sunday, consider choosing 1 of the 7 shirts. We wear any of the 7 shirts. On the second day, the shirt worn on the first day cannot be selected, & we should choose from remaining shirts. The pattern continuous as follows :-

Monday

1. Sunday: 7 shirts to choose from.
2. Monday: 6 shirts to choose from.
3. Tuesday: 5 shirts to choose from.
4. Wednesday: 4 shirts to choose from.
5. Thursday: 3 shirts to choose from.
6. Friday: 2 shirts to choose from.
7. Saturday: 1 shirt to choose from.

• On the last day, saturday, there is only one shirt to choose from, which means, in fact, there is no choice. Put in different names, we are forced on saturday by our choice of which shirt to wear. In this one week, we have to choose one shirt a day, We have six ~~faces~~ free days to choose a shirt. It is the same as saying that our choice of shirt is restricted for one day. So, this week, there are six levels of dfreedom.

08-09-2024

Monday

* Understanding the Degrees of Freedom:-

• An easy way to understand the degrees of mental freedom is by using an example:-

• Consider a sample of data that combines, inorder to simplify, five positive numbers. Values can be any number that does not have a known relationship between them. This data sample theoritically, can have upto five degree of freedom.

• The ~~f~~four numbers in the sample are {3, 8, 5 & 4} & the total number of data samples is expressed as 6; the fifth number should be ~~a~~;

• This should mean that the fifth number should have freedom to It can't be any other. It does not have freedom to be different.

- So, the freedom degrees of this data samples are,
- The free degree formula is equal to the size of sample of data except one.

$\therefore Df = N - 1.$

where as;

Df = Degree of freedom

N = Actual Sample size.

- Degrees of freedom are often discussed in relation to various methods of hypothesis testing in mathematics, such as t-test. It is important to calculate degrees of freedom when trying to understand the importance of t-test arithmetic & the validity of the null hypothesis.

* Degrees of Freedom Formula:-

- The statistical formula to find out how many degrees of freedom are there is quite simple. It implies that degrees of freedom is equivalent to the number of values in a data set minus 1 & appear like this:

$\therefore df = N - 1.$

- Where, N represents the no. of values in data set (sample size).
- That being said, let's have a look at sample calculation.
- If there is a data set of 6, (N=6).
- Call the data set x & make a list with the values for each data.
- For this example data, set x of sample size includes: 10, 30, 15, 25, 45 & 55.

- This data set has a mean or average of 30. Find out mean by adding values & dividing by N:

$(10+30+15+25+45+55)/6 = 30$

- Using the formula, degree of freedom will be computed as $df = N-1$;

- In this example, it appears, $df = 6-1 = 5$.

- This further implies that, in this dataset (sample size), five numbers contain the freedom to vary as long as the mean remains 30.

* Application of Degree of freedom:-

- Although the level of freedom is large & often overlooked concept in mathematics, it is very difficult in the real world.
- For ex, business owners who want to hire employees to produce a product face two changes - function & effect. Additionally, the relationship between employees & output is liability
- In such case, the business owners may determine the amount of product to be produced, which may sufficient for product to be produced. So, in terms of output & staff, owners have one level of freedom.

Q.) Write a difference between one sample t-test & paired t-test.

GoodLuck Page No.
Date

Ans→

one Sample T-Test	paired T-Test
1) This data set is used when, we have one data set for a sample & we need to determine whether this data set belongs to particular population or not. The mean value of population data must be known in this case.	1) The paired t-test is used when we want to compare the means of two selected groups of samples.
2) The formula to determine the T-value in this is, $\therefore t = (\bar{x} - \mu) / (\sigma / \sqrt{n})$	2) For ex, we may use this set to compare the average score of players of an athletics team before & after training program. To calculate this the following formula is used: 1) Calculate sample mean $\therefore \bar{d} = d_1 + d_2 + \dots + d_n$ 2) Calculate sample S.D. $\therefore S = \sqrt{\frac{\Sigma (d - \bar{d})^2}{n-1}}$ 3) Calculate t-value $\therefore t = (\bar{d} / S)\sqrt{n}$

Q.) Compare between T-distribution & Normal distribution with proper diagram?

ns →

T-distribution	Normal distribution
1) T-distribution is defined by it's degree of freedom which itself depends upon the sample size.	1) Normal distribution is defined by it's mean & S.D
2) T-distribution is used when sample size is small.	2) Normal distribution is used when we have large no data points in datasets.
3) It has heavier tail than normal distribution which means more datapoints are away from mean of distribution.	3) It has lighter tail than t-distribution which means more data points lies near mean of population.
4) We use T-distribution in hypothesis testing when S.D of population is unknown	4) Normal distribution is used when S.D is known.
5) T-distribution has a larger range of critical values as compared to normal distribution as this distribution has heavier tail.	5) It has smaller range as compared to t-distribution
6)	6)

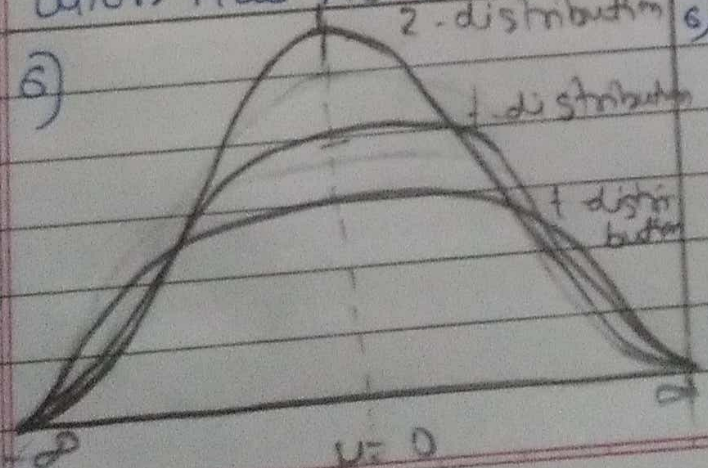


fig t-distribution

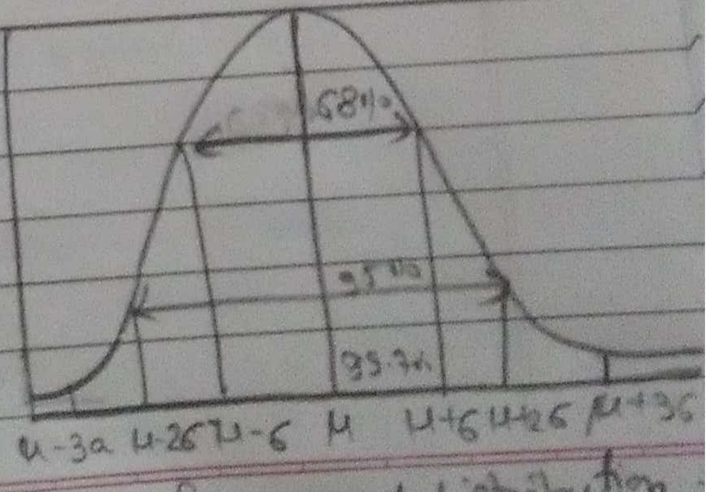


fig-Normal distribution