T-Test:

- The most commonly used test in Statistics is T-Test
- Helps in testing mean (μ).
- Used when the sample size is 30 or ≤ 30 and the population std is unknown.
- T-distribution has been derived mathematically under The assumption of a normally distributed populations.

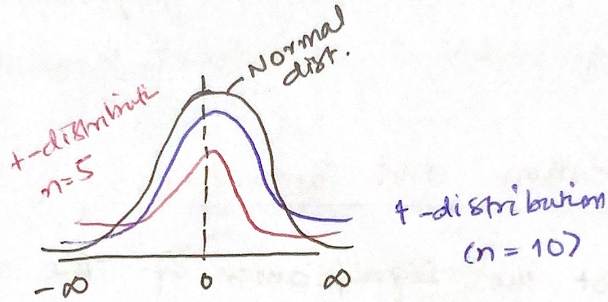
$$f(t) = C\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

where:

C = constant required to make the area under the curve equal to unity.

ν = Degree of freedom. ∴ Independant values in a data sample.

<u>Properties:</u>



1. The variable of t-distribution range from $-\infty$ to ∞ $(-\infty < t < +\infty)$
2. T-distribution will be symmetrical like normal distributio if power of t is even in $f(t)$: pdf (density fn)
3. For large value of ν (ie increased sample size n); the t-distribution tends to a standard normal distribution. This implies that for different ν & shape of t-distrution differs.

sample size	ν	t values	α (0.025) Two tail test
5	4	2.776	
10	9	2.262	$\boxed{z = 1.96}$
30	29	2.045	
∞	–	1.96	Same like t-Test.

4. The t-distribution is less peaked than the normal distribution at the center & higher peaked in tails.

5. The value of y (peak height) attains highest at t_0.

t. Distribution table:

- Gives t values for different level of significance and different degree of freedom.

calculated t value. - compared with tabulated t-value.

- Calculated t-value > tabulated value.
  - significant difference b/w pop. mean & sample mean at 5% level of sign;
- Calc; t-value < tabulated value.
  - No significant difference.

Applications and formulas

1: To test the significance of the mean of a random sample.

$$t = \frac{|\bar{x} - \mu| * \sqrt{n}}{s}, \quad s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

$\bar{x}$: mean of the sample
μ: mean of the population
n: sample size.
s = std of sample.

* we don't know std of the pop.

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{n}}$$

- confidence interval estimates (for α level of significance)
  - one tailed test $\quad \bar{x} \pm t_\alpha \frac{s}{\sqrt{n}}$
  - Two tailed test

$\alpha = 0.05$

To test the difference between means of the two samples (Independent samples): tests are called unpaired

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{S} * \sqrt{\frac{n_1 \times n_2}{n_1 + n_2}} , \quad S = \sqrt{\frac{\Sigma(x_1-\bar{x}_1)^2 + \Sigma(x_2-\bar{x}_2)^2}{n_1+n_2-2}} \rightarrow \nu$$

Combined Std.

$\bar{x}_1$: mean of Sample 1

$\bar{x}_2$: " " " 2

n_1: Sample size of sample 1.

n_2: " " " " 2.

S : Standard deviation (combined)

3. To test the difference between means of two Samples (Dependand samples or matched pair observation)

Paired T-test :

$$t = \frac{\bar{d}\sqrt{n}}{S} , \quad S = \sqrt{\frac{\Sigma(d-\bar{d})^2}{n-1}} \rightarrow \nu = n-1$$

$\bar{d}$: Mean of the differences.

S: Std of the differences.

n: size of the sample.

4. Testing the significance of an observed Correlation coefficient.

$$t = \frac{r}{\sqrt{1-r^2}} * \sqrt{n-2} \rightarrow \nu \text{ (DOF)}$$

r : corr; co-efficient.

n: sample size.

1: Testing the significance of the mean (for random sample)

Q1:

Given data:

- Population mean $\mu = 20$ months.
- life of bulbs in months: 19, 21, 25, 16, 17, 14, 21
- level of significance = 1%

$H_0: \mu = \bar{x}$

$H_a: \mu \neq \bar{x}$

$$t = \frac{|\bar{x} - \mu|}{S} \times \sqrt{n} \; ; \quad S = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n-1}}$$

$\bar{x}$ and S = ?

x	$(x - \bar{x})$	$(x - \bar{x})^2$
19	0	0
21	2	4
25	6	36
16	-3	9
17	-2	4
14	-5	25
21	2	4
$\Sigma x = 133$		$\Sigma (x-\bar{x})^2 = 82$

$\bar{x} = 133/7 = 19$

$S = \sqrt{82/6}$

$S = \sqrt{13.67} \cong 3.7$

$$t = \frac{|19 - 20| * \sqrt{7}}{3.7} = \frac{1 * 2.65}{3.7} = \boxed{0.716}$$

t-value.

Tabulated t-value: ?

$\nu = n - 1 = 6 \qquad t_{0.01} = 3.707$

Ho is passed & accepted. No difference b/w the sample mean & population mean life of a bulb.

- claim is correct

Question No: 02

- Given data:

$\bar{x} = 50$, $\mu = 53$, $n = 15$ & $\Sigma(x-\bar{x})^2 = 130$

$H_0: \mu = \bar{x}$; $H_a: \mu \neq \bar{x}$

$$t = \frac{|\bar{x}-\mu|}{s} * \sqrt{n} \quad ; \quad s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$\Rightarrow s = \sqrt{\frac{130}{14}} = \sqrt{9.29} = 3.05$$

$\therefore \nu = 14$ (DoF)

$$t = \frac{|50-53|}{3.05} * \sqrt{15} = \boxed{3.81}$$

5% level of significance:

tabulated: $t_{0.05} = 2.145$

1% level of significance

$t_{0.01} = 2.977$

95% level of confidence. (sample mean)

$$\bar{x} \pm \frac{s}{\sqrt{n}} * t_{0.05} \Rightarrow 50 \pm \frac{3.05}{\sqrt{15}} * 2.145$$

$$\Rightarrow 50 \pm 1.69$$

limit = 48.31 to 51.69

Sample mean of 99% level of confidence:

$$\bar{x} \pm s/\sqrt{n} * t_{0.01}$$

$\therefore t_{0.01} = 2.977$

$$\Rightarrow 50 \pm \frac{3.05}{\sqrt{15}} * 2.977 = 50 \pm 2.35$$

limit = 47.65 to 52.35

Conclusion:

$\therefore$ H_0 is failed

Significante difference in sample & population mean.

Application #02: Testing difference b/w means (for two Independent samples)

Sol:

$H_0: D_A = D_B$

$H_a: D_A \neq D_B$

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

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

$$\nu = n_1 + n_2 - 2$$

Given data:

- D_A (in kg): 8, 10, 12, 9, 14, 13
- D_B (in kg): 7, 9, 14, 12, 8

$n_1 = 6$, $n_2 = 5$

Calc:

x_1	$x_1 - \bar{x}_1$	$(x_1 - \bar{x}_1)^2$	x_2	$x_2 - \bar{x}_2$	$(x_2 - \bar{x}_2)^2$
8	-3	9	7	-3	9
10	-1	1	9	-1	1
12	1	1	14	4	16
9	-2	4	12	2	4
14	3	9	8	-2	4
13	2	4	-		
$\Sigma = 66$		$\Sigma = 28$	$\Sigma = 50$		$\Sigma = 34$

$\bar{x}_1 = 66/6 = 11$

$\bar{x}_2 = \frac{50}{5} = 10$

$$S = \sqrt{\frac{28+34}{9}} = \sqrt{66/9} = \sqrt{6.88} = 2.62$$

$$t = \frac{|11 - 10|}{2.62} * \sqrt{\frac{30}{11}} = 1/2.62 * \sqrt{30/11} = 1/2.62 * \sqrt{2.72}$$

$t = 0.63$ (calculated t-value)

$\nu = 9$

$t_{0.05} = 2.262$

H_0 is passed i.e., $Drug_A = Drug_B$. Both have the same efficacy.

QUESTION NO: 02

Solution:

H_0 : $Lab_A = Lab_B$ $\qquad$ H_a: $Lab_A \neq Lab_B$

Given data: (Sugar content in chocolate in mg/g)

Lab_A : 8, 9, 9, 6, 4, 6

Lab_B : 7, 8, 6, 4, 5, 6

$n_1 = 6$, $n_2 = 6$

x_1	$(x_1-\bar{x}_1)$	$(x_1-\bar{x}_1)^2$	x_2	$(x_2-\bar{x}_2)$	$(x_2-\bar{x}_2)^2$
8	1	1	7	+1	1
9	2	4	8	2	4
9	2	4	6	0	0
6	-1	1	4	-2	4
4	-3	9	5	-1	1
6	-1	1	6	0	0
$\Sigma = 42$		$\Sigma = 20$	$\Sigma x_2 = 36$		$\Sigma = 10$
$\bar{x}_1 = 42/6 = 7$			$\bar{x}_2 = 36/6 = 6$		

$$S = \sqrt{\frac{20+10}{10}} = \sqrt{30/10} = \sqrt{3} = 1.732$$

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

$$\Rightarrow \frac{|7-6|}{1.732} * \sqrt{36/12} = \frac{1}{1.732} \times 1.732 = 1$$

(calculated) t-value

$\nu = 10$

$t_{0.05} = 2.228$ (tabulated)

H_0 is passed & accepted. So, there is no significant difference b/w the sugar content obtained by Lab_A & Lab_B

Application # 03:

Testing of difference between means (for two dependent samples)

Sol: Given data:

Blood pressure before Drug A = 112, 113, 118, 120, 119, 113, 110, 122

Blood pressure after Drug A = 116, 120, 117, 125, 126, 111, 111, 117.

H_0: BP before Drug A = BP after Drug A

H_a: BP before Drug A $\neq$ BP after Drug A

$$t = \frac{\bar{d} \times \sqrt{n}}{S} \quad ; \quad S = \sqrt{\frac{\Sigma(d-\bar{d})^2}{n-1}}$$

Calc $\bar{d}$ & S.

Before D_A	After D_A	d	$(d-\bar{d})$	$(d-\bar{d})^2$
112	116	4	2	4
113	120	7	5	25
118	117	-1	-3	9
120	125	5	3	9
119	126	7	5	25
113	111	-2	-4	16
110	111	1	-1	1
122	117	-5	-7	49

$\Sigma d = 16$

$\bar{d} = 16/8 = 2$

$\Sigma(d-\bar{d})^2 = 138$

$$S = \sqrt{138/7} = 4.44$$

$$t = \frac{2 \times \sqrt{8}}{4.44} = 1.912$$ Calculated t-value

$\nu = 7$

(tabulated) $t_{0.05} = 2.365$

H_0 is correct & accepted. So, the drug has no significant role in the change of the blood pressure.

Testing the Significance
(for observed Coefficient)

Solution:

H_0 : Correlation = Not significant.

H_a: Correlation ≠ not significant.

Given data:

$n = 27$, $r = 0.55$

$$t = \frac{r}{\sqrt{1-r^2}} \times \sqrt{n-2}$$

r = correlation coefficient

n = sample size

$$t = \frac{0.55}{\sqrt{1-(0.55)^2}} \times \sqrt{27-2}$$

$$t = \frac{0.55}{\sqrt{1-0.3025}} \times \sqrt{25}$$

$$t = 0.55 / \sqrt{0.6975} \times 5$$

$t = \boxed{3.293}$ (calculated value)

$v = n-2 \Rightarrow 25$

$t_{0.05}$ (tabulated value) = 2.060

H_0 is failed & rejected. So, variable in the population is non-correlated.

Q#02:

Given data:

$n = 11$, $r = 0.6$

$$t = \frac{r}{\sqrt{1-r^2}} \times \sqrt{n-2}$$

$$t = \frac{0.6}{\sqrt{1-(0.6)^2}} \times \sqrt{11-2}$$

$$t = \frac{0.6}{\sqrt{1-0.36}} \times \sqrt{9}$$

$$t = \frac{0.6}{\sqrt{0.64}} \times 3 = \boxed{2.25} \text{ (calculated value)}$$

$v = n-2$

$t_{0.05} = \boxed{2.262}$ (Tabulated Value)

* H_0 is Correct & passed so, correlation coefficient is not significant.