Tests of Significance For Small Samples

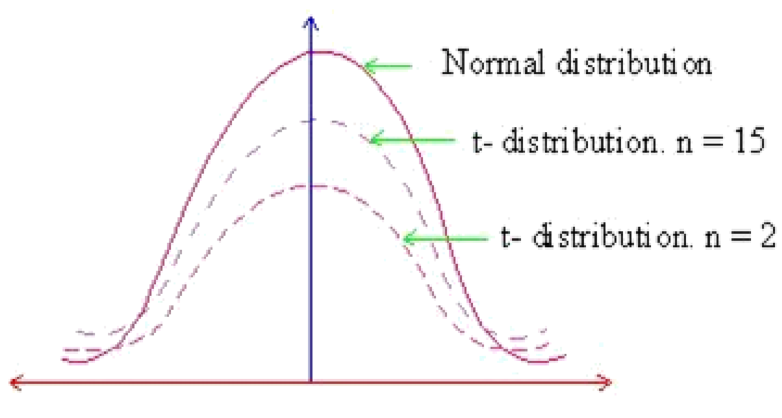
6.1 Introduction

For small samples (size < 30); tests proposed for large samples do not hold good as sampling distribution cannot be assumed to be normal for small samples. This led to search of new approaches to deal with small samples. It should be made rational that the methods and theory applicable to small samples can be used for large samples; but the converse is not true

After hypothesis formulation; choice of test may be sometimes baffling unless specified which test to use. Following suggestions should be kept in mind while choosing test of significance for any hypothesis.

- Size of sample: If size of sample is greater than thirty, use any of the applicable large sample tests or Chi-Square test depending upon the applicability.
- Variance of population: If the population variance is known and the underlying distribution is normal, z-test should be used. Also with known variance; if the distribution is not normal, yet for large sample size, z test can be used. But in case of unknown variance,
  t-test should be used.

Degree of freedom (ν): The degrees of freedom are the number of independent quantities that can be assigned to any statistical distribution arbitrarily. Suppose we are required to choose 5 numbers whose sum is 30, then we have choice of only 4 numbers. This suggests that a data of size n has $(n-1)$ degrees of freedom in general. In case of two restrictions, degrees of freedom will be $(n-2)$.



6.2 Student's t-Distribution

Gosset; who wrote under the pen-name 'Student', derived a theoretical distribution known as Student's t distribution; which is used to test a hypothesis when the sample size is small and the population variance is not known. As n (size of the sample) increases; $t$distribution tends to normal distribution. It is evident from the graph below that t-distribution has proportionally larger area at its tails than the normal distribution.

If $x_1, x_2, \cdots, x_n$ be a random sample of size $n \leq 30$, from a normal distribution with mean μ and variance σ^2(not known); also $\overline{x}$ be the sample mean and s standard deviation of the sample

then $t = \frac{\overline{x}-\mu}{\frac{s}{\sqrt{n}}}$, where $\overline{x} = \frac{\sum x}{n}$, $s = \sqrt{\frac{\sum(x-\overline{x})^2}{n-1}}$

Degree of freedom is given by $\nu = n - 1$

Statistics t is used to test the significance of:

1. Sample mean
2. Difference between two sample means
3. Sample coefficient of correlation
4. Sample Regression Coefficient

Result *I* Testing significance of sample mean

Set up the hypothesis H_0: sample is drawn from the given population.

To calculate significance of sample mean at 5% level, calculate the statistics

$t = \frac{\overline{x}-\mu}{\frac{s}{\sqrt{n}}}$ and compare with the table value of t at $(n-1)$ degrees of freedom.

Let the tabulated value of t be t_1, if $|t| < t_1$; H_0, is accepted, otherwise value of t is considered significant and we reject the hypothesis H_0.

Fiducial (Confidence) limits of population mean

At 5% significance level (95% confidence level):

$\overline{x} \pm \frac{s}{\sqrt{n}} t_{0.05}$, where $t_{0.05}$ is table value of t at given degrees of freedom

At 1% significance level (99% confidence level):

$\overline{x} \pm \frac{s}{\sqrt{n}} t_{0.01}$, where $t_{0.01}$ is table value of t at given degrees of freedom.

Example1 The mean life time of sample of 16 fluorescent light bulbs produced by a company is computed to be 1550 hours with a standard deviation of 100 hours. The company claims that the average life of the each bulb is 1600 hours. Using the level of significance of 0.05, is the claim accepted? Also find the confidence limits for μ.

Solution: Let H_0: Average life of the each bulb is 1600 hours.

Population mean (μ) = 1600 hours, sample size (n) = 16

Sample mean $\overline{x} = 1550$, sample standard deviation (s) = 100 hours

$$\text{Statistic } t = \frac{\overline{x}-\mu}{s}\sqrt{n} = \frac{1550-1600}{100}\ (4) = -2$$

Table value of t at 5% level of significance for 15 degrees of freedom is 2.13.

∴ Calculated value of $|t| <$ table value of t, hence the hypothesis (H_0) is accepted that the average life of the each bulb is 1600 hours.

Also confidence limits for $\mu = \overline{x} \pm \frac{s}{\sqrt{n}}\ t_{0.05} = 1550 \pm \frac{100}{\sqrt{16}}\ (2.13)$

$$=1496.75 \text{ to } 1603.25$$

Example2 Ten individuals are chosen at random from a population and their heights are found in inches as: 63, 63, 64, 65, 66, 69, 69, 70, 70, and 71. Discuss the suggestion that the mean height of universe is 65. Value of t at 5% level of significance for 9 degrees of freedom is 2.262.

Solution: Let H_0: The sample is drawn from the given population.

Population mean (μ) = 65, sample size (n) =10

Calculating sample mean and standard deviation:

x	$x-\overline{x}$, $\overline{x}=67$	$(x-\overline{x})^2$
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
$\sum x = 670$ $\overline{x} = \frac{670}{10} = 67$		$\sum(x-\overline{x})^2 = 88$

Sample standard deviation: $s = \sqrt{\frac{(x-\overline{x})^2}{n-1}} = \sqrt{\frac{88}{9}} = 3.13$

Statistic $t = \frac{\bar{x}-\mu}{s}\sqrt{n} = \frac{(67-65)\sqrt{10}}{3.13} = \frac{2\sqrt{10}}{3.13} = 2.02$

$\therefore$ Calculated value of t is less than tabulated value at 9 degrees of freedom.

Hence the hypothesis (H_0) is accepted that sample is drawn from the given population i.e. mean height of the universe is 65 inches.

Result *II* Testing difference between means of two small samples

Let two independent samples $(x_1, x_2, \cdots, x_{n_1})$; $(y_1, y_2, \cdots, y_{n_2})$ of sizes n_1 ; n_2 with means

$\bar{x}$; $\bar{y}$ and standard deviations $s_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n_1-1}}$; $s_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n_2-1}}$ be drawn from two normal populations with means μ_1 ; μ_2 and having same variance σ^2, then to test whether the two population means μ_1 and μ_2 are same:

Calculate $t = \frac{\bar{x}-\bar{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$; where $s = \sqrt{\frac{\sum(x-\bar{x})^2+\sum(y-\bar{y})^2}{n_1+n_2-2}}$ or $\sqrt{\frac{(n_1-1)s_x^2+(n_2-1)s_y^2}{n_1+n_2-2}}$

and compare with table value of t (say t_1) at (n_1+n_2-2) degrees of freedom. Hypothesis is accepted if calculated value of $|t| < t_1$.

Example3 Two independent samples showing weights in ounces of eight and seven items are given below:

Sample I:	10	12	13	11	15	9	12	14
Sample II:	9	11	10	13	9	8	10	

Is the difference between the means of the samples significant?

Value of t at 5% level of significance for 15 degrees of freedom is 2.16.

Solution: Let the null hypothesis (H_0) be: $\mu_1 \approx \mu_2$

Here $n_1 = 8$ and $n_2 = 7$, calculating sample means and standard deviations

x	$(x-\bar{x})$	$(x-\bar{x})^2$	y	$(y-\bar{y})$	$(y-\bar{y})^2$
10	-2	4	9	-1	1
12	0	0	11	1	1
13	1	1	10	0	0
11	-1	1	13	3	9
15	3	9	9	-1	1
9	-3	9	8	-2	4

12	0	0	10	0	0
14	2	4			
$\sum x = 96$ $\overline{x} = \frac{96}{8} = 12$		28	$\sum y = 70$ $\overline{y} = \frac{70}{7} = 10$		16

$$s = \sqrt{\frac{\sum(x-\overline{x})^2+\sum(y-\overline{y})^2}{n_1+n_2-2}} = \sqrt{\frac{\sum(x-12)^2+\sum(y-10)^2}{8+7-2}} \sqrt{\frac{28+16}{8+7-2}} = 1.8$$

$$\therefore t = \frac{\overline{x}-\overline{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{12-10}{(1.8)\sqrt{\frac{1}{8}+\frac{1}{7}}} = 2.15$$

Table value of t at 5% level of significance for 13 degrees of freedom is 2.16

$\therefore$ Calculated value of $|t| <$ table value of t

Hence the hypothesis (H_0) is accepted that $\mu_1 \approx \mu_2$, i.e. difference between the means of the two samples is not significant.

Example4 Two types of batteries were tested for their mean life length and following results are obtained. Is there a significant difference in the two batteries?

	Sample size	Mean life	Variance
Battery A	10	500 hours	100
Battery B	8	540 hours	81

Solution: Let the null hypothesis (H_0) be: $\mu_1 \approx \mu_2$ i.e. there is no significant difference in the two batteries. Here $n_1 = 10$ and $n_2 = 8$, $\overline{x} = 500$ and $\overline{y} = 540$, $s_x = 10$, $s_y = 9$

$$\Rightarrow s = \sqrt{\frac{(n_1-1)s_x^2+(n_2-1)s_y^2}{n_1+n_2-2}} = \sqrt{\frac{9(100)+7(81)}{10+8-2}} = 9.58$$

$$\therefore t = \frac{\overline{x}-\overline{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{500-540}{(9.58)\sqrt{\frac{1}{10}+\frac{1}{8}}} = -8.8$$

Table value of t at 5% level of significance for 16 degrees of freedom is 2.12.

$\therefore$ Calculated value of $|t| \gg$ table value of t, hence the hypothesis (H_0) is rejected that $\mu_1 \approx \mu_2$, i.e. difference between two batteries is highly significant.

Example5: A set of 15 observations has mean 68.57; standard deviation 2.4, another set of 7 observations gives mean as 64.14; standard deviation 2.7.

Use t-test to find whether two sets of data are drawn from populations with same mean, if standard deviations of two populations are assumed to be equal.

Solution: Let the null hypothesis (H_0) be: $\mu_1 \approx \mu_2$

i.e. there is no significant difference between the two population means.

Here $n_1 = 15$ and $n_2 = 7$, $\overline{x} = 68.57$ and $\overline{y} = 64.14$

$s_x = 2.4,\ s_y = 2.7$

$$\Rightarrow s = \sqrt{\frac{(n_1-1)s_x^2 + (n_2-1)s_y^2}{n_1+n_2-2}} = \sqrt{\frac{14(2.4)^2+6(2.7)^2}{15+7-2}} = 2.49$$

$$\therefore\ t = \frac{\overline{x}-\overline{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{68.57-64.14}{(2.49)\sqrt{\frac{1}{15}+\frac{1}{7}}} = 3.89$$

Table value of t at 5% level of significance for 20 degrees of freedom is 2.086

$\therefore$ Calculated value of $|t| >$ table value of t, hence the hypothesis (H_0) is rejected that $\mu_1 \approx \mu_2$, i.e. difference between two means is significant.

Result *III* Testing difference between two dependent samples or paired observations

A 't' test can be efficiently used to compare two samples of the same population for some treatment effects; for instance to compare efficacy of two drugs on a population or to study the effect of coaching on some students etc.

Compute the statistic $t = \frac{\overline{x}-\Delta}{\frac{s}{\sqrt{n}}}$; where Δ denotes targeted value and is zero for testing equal means and $s = \sqrt{\frac{\sum d^2}{n-1}}$, where d is the deviation from the mean difference.

Example6 A dietitian opts to try out a new type of diet program on ten overweight girls for 2 months. He targets to make them loose 6 kgs on average, and records their weights before and after the diet program. Use 0.05 significance level to test whether this special diet program helped or not.

Before weights	65	77	99	86	84	93	59	72	69	103
After weights	63	72	92	80	80	87	57	67	64	95

Table value of t at 5% level of significance for 9 degrees of freedom is 2.26.

Solution: Let H_0: Average weight loss caused by diet program is 6 kg.

Calculating deviations from the mean weight loss:

Girls	Weight difference (m)	$d = m - \overline{x}$	d^2
1	-2	+3	9
2	-5	0	0
3	-7	-2	4
4	-6	-1	1
5	-4	+1	1
6	-6	-1	1
7	-2	+3	9
8	-5	0	0
9	-5	0	0
10	-8	-3	9
$n = 10$	$\sum m = -50$, $\overline{x} = \frac{\sum m}{n} = \frac{-50}{10} = -5$		$\sum d^2 = 34$

Standard deviation of the differences $s = \sqrt{\frac{\sum d^2}{n-1}} = \sqrt{\frac{34}{9}} = 1.94$

Statistic $t = \frac{\overline{x}-\Delta}{\frac{s}{\sqrt{n}}}$, where targeted weight loss $(\Delta) = -6$ kg

$\Rightarrow t = \frac{-5-(-6)\sqrt{10}}{1.94} = 1.63$ $\therefore$ Calculated value of $|t| <$ table value of t

Hence the hypothesis (H_0) is accepted that the average weight loss caused by diet program is 6 kg.

Example7 Two laboratories carried out independent estimates of lead content (in mg) in noodles of a certain brand. A sample is taken from each batch, halved and the separate halves were tested in the two laboratories to obtain the following results:

Batch Number	1	2	3	4	5	6	7	8	9	10
Lab A	9	8	8	4	7	7	9	6	6	6
Lab B	7	8	7	3	8	6	9	4	7	8

Does the testing suggest same average lead content in the brand?

Solution: H_0: Average lead content of two samples are equal.

Calculating deviations from the mean difference:

Batch	Difference (m)	$d = m - \overline{x}$	d^2
1	2	1.7	2.89
2	0	-0.3	0.09
3	1	0.7	0.49
4	1	0.7	0.49
5	-1	-1.3	1.69
6	1	0.7	0.49

7	0	-0.3	0.09
8	2	1.7	2.89
9	-1	-1.3	1.69
10	-2	-2.3	5.29
$n = 10$	$\sum m = 3$, $\bar{x} = \frac{\sum m}{n} = \frac{3}{10} = 0.3$		$\sum d^2 = 16.10$

Standard deviation of the differences $s = \sqrt{\frac{\sum d^2}{n-1}} = \sqrt{\frac{16.1}{9}} = 1.34$

$t = \frac{\bar{x}-\Delta}{\frac{s}{\sqrt{n}}} = \frac{0.3-(0)\sqrt{10}}{1.34} = 0.71$ $\because \Delta = 0$ as we are testing for equal means of two lab tests

Table value of t at 5% level of significance for 9 degrees of freedom is 2.26.

$\therefore$ Calculated value of $|t| <$ table value of t at 9 degrees of freedom at 5% level of significance, hence the hypothesis (H_0) is accepted that the average lead content of two samples are equal.

Result IV Testing significance of population correlation coefficient from sample coefficient of correlation

Let(x_1, y_1), (x_2, y_2), ... , (x_n, y_n) having coefficient of correlation r, be a random sample from a bivariate frequency distribution having individual means μ_1 ; μ_2 and standard deviations σ_1; σ_2. Then to test the hypothesis that population correlation coefficient (ρ) is zero: Compute the statistic $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$, where n is the sample size

Hypothesis is accepted if calculated value of $|t|$ is less than tabulated value of t at

$(n-2)$degrees of freedom for the specified significance level.

Example8 A random sample of size 18 from a bivariate population gave correlation coefficient $r = 0.4$. Does this indicate the existence of correlation in the population?

Solution: Let H_0: Population correlation coefficient (ρ) is zero, i.e. there is no correlation between the population variables.

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{(0.4)\sqrt{18-2}}{\sqrt{1-(0.4)^2}} = 1.7$$

Table value of t at 16 degrees of freedom is 2.12

$\therefore$ Calculated value of $|t|$ is less than tabulated value of t at 16 degrees of freedom, hence the hypothesis H_0 is accepted that there is no correlation between the population variables.

6.3 Snedecor's F – test for Testing Equality of Two Population Variances

An F -test is used to test if the variances of two populations are equal. It can be a one-tailed or two-tailed test. The one-tailed version tests in only one direction, i.e. variance of the first population is either greater or less than the second population but not both ways. The two-tailed version tests for the hypothesis that the variances are not equal but one can be greater or less than the other.

Let two independent random samples ($x_1, x_2, \cdots, x_{n_1}$) ; ($y_1, y_2, \cdots, y_{n_2}$) having standard deviations $s_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n_1-1}}$; $s_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n_2-1}}$ be drawn from two normal populations. Snedecor defined the statistic $\boldsymbol{F} = \frac{s_x^2}{s_y^2}$, $s_x^2 > s_y^2$ for testing equality of two population variances. Here greater of the two variances s_x^2 and s_y^2 is to be taken in the numerator and if n_1 corresponds to the greater variance, then degree of freedom is $(n_1 - 1, n_2 - 1)$.

If calculated value of F is less than the table value of F with $\nu = (n_1 - 1, n_2 - 1)$ at given level of significance, the null hypothesis that 'the two samples might have been drawn from two normal population with the same variance' is accepted.

Example9 Two samples of size 9 and 8 give the sum of squares of deviations from their respective means equal to 160 inches square and 91 inches square respectively. Can they be regarded as drawn from two normal populations with same variances? (Value of F for 8 and 7 degrees of freedom is 3.73)

Solution: Let H_0: Two samples have been drawn from two normal populations with same variance.

$$\text{Given } s_x^2 = \frac{\sum(x-\bar{x})^2}{n_1-1} = \frac{160}{9-1} = 20, \quad s_y^2 = \frac{\sum(y-\bar{y})^2}{n_2-1} = \frac{91}{8-1} = 13$$

$$\therefore \boldsymbol{F} = \frac{s_x^2}{s_y^2} = \frac{20}{13} = 1.54 \qquad \because s_x^2 > s_y^2$$

Table value of $\boldsymbol{F}$ for (8, 7) degrees of freedom is 3.73, $\therefore$ Calculated value of F is much less than the table value of $\boldsymbol{F}$ at (8,7) degrees of freedom. Hence H_0 is accepted, i.e. two samples may be regarded as drawn from two normal populations with same variances.

Example10 Show how we can use Student's- $\boldsymbol{t}$ test and Snedecor's $\boldsymbol{F}$ test to decide whether the following two samples have been drawn from the same normal population. Which of the two tests would you apply first and why?

	Size	Mean	Sum of squares of deviations from the mean

Sample I	9	68	36
Sample II	10	69	42

Given that $t_{0.5}(17) = 2.11$, $F_{0.5}(9{,}8) = 3.39$

Solution: H_0: Two samples have been drawn from the same normal population

To test H_0 using Student's- $\boldsymbol{t}$ test, population variance σ^2 should be same, we can test that two samples have been drawn from two normal populations with same variance using Snedecor's $\boldsymbol{F}$ test. $\therefore$ Snedecor's $\boldsymbol{F}$ test should be applied first.

For the two samples I and II say ($x_1, x_2, \cdots, x_9$) ; ($y_1, y_2, \cdots, y_{10}$)

$$s_x^2 = \frac{\sum(x-\bar{x})^2}{n_1-1} = \frac{36}{9-1} = 4.5, \quad s_y^2 = \frac{\sum(y-\bar{y})^2}{n_2-1} = \frac{42}{10-1} = 4.67$$

$$\therefore \boldsymbol{F} = \frac{s_y^2}{s_x^2} = \frac{4.67}{4.5} = 1.04 \quad \because s_y^2 > s_x^2$$

Table value of $\boldsymbol{F}$ for (9, 8) degrees of freedom is 3.39

$\therefore$ Calculated value of F is much less than the table value of F at (9, 8) degrees of freedom.

Hence the two samples may be regarded as drawn from two normal populations with same variances.

Again to test whether the two population means are same using t-test

Calculate $t = \dfrac{\bar{x}-\bar{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$; where $s = \sqrt{\dfrac{(n_1-1)s_x^2 + (n_2-1)s_y^2}{n_1+n_2-2}}$

$$\Rightarrow s = \sqrt{\frac{(9-1)(4.5)+(10-1)(4.67)}{9+10-2}} = 2.14$$

$$\therefore t = \frac{68-69}{(2.14)\sqrt{\frac{1}{9}+\frac{1}{10}}} = -1.02$$

Table value of t at 5% level of significance for 17 degrees of freedom is 2.11.

$\therefore$ Calculated value of $|t| <$ table value of t, hence the hypothesis that two population means are same is accepted.

Thus using Snedecor's F test, we can say that the two samples have been drawn from two normal populations with same variance and also the two population means are same using t-test, hence we can conclude that the hypothesis H_0 that the two samples have been drawn from the same normal population may be accepted.

6.4 Fisher's Z Test for Testing Significance of Correlation Coefficient for Small Samples

If r is the correlation co-efficient of a sample and ρ be the population co-efficient of correlation, then to test the hypothesis that the given sample has been drawn from the population whose coefficient of correlation is ρ:

Compute the statistic $\frac{z-\xi}{\frac{1}{\sqrt{n-3}}}$, where $z = \frac{1}{2}\log_e\left(\frac{1+r}{1-r}\right)$ and $\xi = \frac{1}{2}\log_e\left(\frac{1+\rho}{1-\rho}\right)$

If value of $\left|\frac{z-\xi}{\frac{1}{\sqrt{n-3}}}\right| < 1.96$, hypothesis is accepted at 5% level of significance.

Example11 Test the significance of the correlation $r = 0.6$ for a sample of size 20 against hypothetical population coefficient of correlation $\rho = 0.8$

Solution: Let H_0: correlation coefficient of population is 0.8

$$\text{Here } z = \frac{1}{2}\log_e\left(\frac{1+r}{1-r}\right) = \frac{1}{2}\log_e\left(\frac{1+0.6}{1-0.6}\right) = 0.693$$

$$\xi = \frac{1}{2}\log_e\left(\frac{1+\rho}{1-\rho}\right) = \frac{1}{2}\log_e\left(\frac{1+0.8}{1-0.8}\right) = 1.099$$

$$\text{Now } \frac{z-\xi}{\frac{1}{\sqrt{n-3}}} = \frac{0.693-1.099}{\frac{1}{\sqrt{20-3}}} = (0.693 - 1.099)\sqrt{17} = -1.67$$

$$\therefore \left|\frac{z-\xi}{\frac{1}{\sqrt{n-3}}}\right| = 1.67 < 1.96$$

Hence the sample may be regarded as coming from population with coefficient of correlation $(\rho) = 0.8$

Exercise 6

1 A factory makes a machine part with axle diameter of 0.7 inch. A random sample of 10 parts shows a mean diameter of 0.742 inch with a standard deviation of 0.04 inch. On the basis of this sample would you say that the work is inferior?

Value of t at 5% level of significance for 9 degrees of freedom is 2.262.

2 A random sample of 10 boys had the I.Q. levels: 70, 120, 110, 101, 88, 83, 95, 98, 107 and 100. Does this data support the assumption of population mean I.Q. of 100 at 5% level of significance?

3 In a school the heights of six randomly chosen girls are: 63, 65, 68, 69, 71 and 72 inches and those of nine randomly chosen boys are 61, 62, 65, 66, 69, 70, 71, 72 and 73 inches. Discuss the hypothesis that the girls are taller than boys.

Value of t at 5% level of significance for 13 degrees of freedom is 1.77.

4 A random sample of size 16 has mean 53. The sum of squares of deviations from the mean is 135. Can the sample be regarded as taken from a population having mean as 56?

5 A new medicine is given to 12 patients whose B.P. increases by 5, 2, 8, -1, 3, 0, -2, 1, 5, 0, 4, 6 units. Can we conclude that the medicine results in increased blood pressure?

6 Two samples of different brands were tested for average life; a sample from first brand of size 7 shows a mean life of 1036 hours with a standard deviation of 40 hours and a sample of size 8 shows a mean life of 1234 hours with a standard deviation of 36 hours. Is the difference in the two sample means significant to conclude that the second brand has more life than first brand?

7 A researcher hypothesizes that people who are allowed to sleep for only four hours will score significantly lower in an objective skills test than people who are allowed to sleep for eight hours. He selects sixteen participants and randomly assigns them to one of two groups. In one group he makes participants sleep for eight hours and in the other group he allows them to sleep only for four hours. The next morning he administers the skill test to all participants. Scores range from 1-9 with high scores representing better performance.

Test scores→								
8 hours sleep group	5	7	5	3	5	3	3	9
4 hours sleep group	8	1	4	6	6	4	1	2

Test the hypothesis, given that $t_{0.5}(14) = 2.145$

8 A group of 10 rats fed on the diet A and another group of 8 rats fed on the diet B recorded the following increase in weights in a week:

Weight gains (grams)→										
Diet A	5	6	8	1	12	4	3	9	6	10
Diet B	2	3	6	8	10	1	2	8	-	-

Does it show superiority of Diet A over that of Diet B?

9 Test runs with 6 models of an experimental engine showed that they operated for 24, 28, 21, 23, 32 and 22 minutes with a gallon of fuel. If the probability of a Type I error is at most 0.01, is this an evidence against the hypothesis that on average this kind of engine will operate for at least 27 minutes per gallon on the same fuel?

10 Test whether the following two samples have been drawn from the same normal population.

	Size	Mean	Sum of squares of deviations from the mean
Sample I	10	15	90
Sample II	12	14	108

Given that $t_{0.5}(20) = 2.086$, $F_{0.5}(9,11) = 2.9$

Answers

1. $t = 3.16$ which is greater than the table value at 5% significance level, $\therefore$ work can be considered to be inferior.
2. $t = 0.62$ $\therefore$ given data supports the mean I.Q. as 100.
3. $t = 0.3031$ $\therefore$ there is no significant difference between the sample means
4. No
5. No
6. $t = 18.15$ $\therefore$ significantly different to conclude that the second brand has more life than first brand
7. $t = 0.847$ $\therefore$ there is no significant difference between the performances of two groups.
8. No
9. No
10. Yes, the two samples can be considered to be drawn from the same normal population.