1) Conduct f-test for the following samples and use alpha = 0.025 :
sampel -1 having sample size= 41 and variance =109.63
sampel -2 having sample size =21 and variance =65.99
Topic : ANOVA

formula:
F = s1^2 / s2^2

Sample 1: n1 = 41 s1^2 = 109.63
Sample 2: n2 = 21 s2^2 = 65.99
Calculating the F-statistic:
F = s1^2 / s2^2 = 109.63 / 65.99 = 1.660
Calculating the degrees of freedom:
df1 = n1 - 1 = 40 df2 = n2 - 1 = 20

Using an F-distribution table with a significance level of 0.025 and degrees of freedom 40 and 20, we find the critical value to be 2.284.

variances of sample 1 and sample 2 are significantly different at a significance level of 0.025.

2) What is analysis of variance (ANOVA) and where it is used? Give the mathematical model of one way ANOVA.
Topic: Hypothesis testing for small samples (t- distribution)

Analysis of variance (ANOVA) is a statistical technique used to determine if there is a significant difference between the means of two or more groups. It compares the variance between groups to the variance within groups to determine if there is a significant difference in means.

ANOVA is used in many fields including psychology, social sciences, economics, biology, and engineering. It can be used to compare means in experiments where there are multiple groups, such as in clinical trials, and to identify which group or groups are significantly different from the others.

The mathematical model of one-way ANOVA is:
$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$

To test the null hypothesis, we calculate the F-statistic using the formula:
F = MS_between / MS_within

To conduct hypothesis testing using the t-distribution, we calculate the t-statistic using the formula:

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

where $\bar{x}$ is the sample mean, μ is the hypothesized population mean, s is the sample standard deviation, and n is the sample size.
We then use a t-distribution table or statistical software to find the critical value for the t-statistic at the chosen significance level and degrees of freedom.

3) Retail sales record shows that average monthly expenditure per family for a certain commodity was Rs. 255. A random sample of say10 families showed the mean expenditure to be Rs. 285 with standard deviation of Rs. 40. Is there any significant change in the average recorded earlier ?

alpha, $\alpha = 0.05$ - two tailed test

$H0 : \mu = 255$
$H1 : \mu \neq 255$
$to = \bar{x} - \mu / (\sigma/\surd n) = 285 - 255 / (40/\surd 10) = 2.37$
DoF = n - 1 = 9
te = 2.26
Inference: to > te is highly significant change in the average recorded earlier. Hence H0 rejected.

4) State the situation in which paired t-tests are used.
Topic: Hypothesis Testing for large sample size (Z-test for one sample and two samples)

We calculate the t-statistic using the formula:
$t = (\bar{x}d - \mu d) / (sd / \surd n)$

where $\bar{x}d$ is the sample mean of the differences, μd is the hypothesized population mean of the differences, sd is the sample standard deviation of the differences, and n is the sample size.

The formula for the z-statistic for a one-sample test is:
$z = (\bar{x} - \mu) / (\sigma / \surd n)$

The formula for the z-statistic for a two-sample test is:
$z = (\bar{x}1 - \bar{x}2) / (\sigma / \surd(n1 + n2))$

where $\bar{x}1$ and $\bar{x}2$ are the sample means for the two groups, σ is the population standard deviation (assumed to be equal for both groups), and n1 and n2 are the sample sizes for the two groups.

5) The mean diastolic blood pressure for a group of 81 adults was found to be 79.2 mm. Test the hypothesis that the mean diastolic blood pressure is 75 mm at 5% level of significance and give conclusion. Population standard deviation is known to be 9 mm

The null hypothesis (H0) is that the mean diastolic blood pressure is 75 mm, and the alternative hypothesis (Ha) is that the mean diastolic blood pressure is not 75 mm.

We can use the following formula to calculate the test statistic:
$z = (\bar{x} - \mu) / (\sigma / \sqrt{n})$

where: $\bar{x}$ = sample mean = 79.2 mm μ = hypothesized population mean = 75 mm σ = population standard deviation = 9 mm n = sample size = 81

Plugging in the values, we get:
$z = (79.2 - 75) / (9 / \sqrt{81})$ $z = 4.2 / 1$ $z = 4.2$

The calculated z-value of 4.2 is the test statistic. We will now compare this to the critical value from the standard normal distribution table at 5% level of significance.

At 5% level of significance, the critical value for a two-tailed test is ±1.96. Since the calculated z-value of 4.2 is greater than the critical value of 1.96, we reject the null hypothesis.

Conclusion: There is sufficient evidence to conclude that the mean diastolic blood pressure for the population is not 75 mm at 5% level of significance.

6) The sample of 256 bricks has a mean weight of 2.12 kg with standard deviation 560 gm. Test the hypothesis that sample come from the population with a mean weight of 2 kg at 5% level of significance.

The null hypothesis (H0) is that the mean weight of the population is 2 kg, and the alternative hypothesis (Ha) is that the mean weight of the population is not 2 kg.

We can use the following formula to calculate the test statistic:
$t = (\bar{x} - \mu) / (s / \sqrt{n})$
where: $\bar{x}$ = sample mean = 2.12 kg μ = hypothesized population mean = 2 kg s = sample standard deviation = 560 gm = 0.56 kg n = sample size = 256

Plugging in the values, we get:
$t = (2.12 - 2) / (0.56 / \sqrt{256})$ $t = 0.12 / 0.035$ $t = 3.43$
The calculated t-value of 3.43 is the test statistic. We will now compare this to the critical value from the t-distribution table at 5% level of significance with 255 degrees of freedom (df = n - 1).

At 5% level of significance with 255 degrees of freedom, the critical values for a two-tailed test are ±1.96. Since the calculated t-value of 3.43 is greater than the critical value of 1.96, we reject the null hypothesis

Conclusion: There is sufficient evidence to conclude that the sample does not come from a population with a mean weight of 2 kg at 5% level of significance.

7) In a college there are two faculties Arts and Science. The average weight of students in the sample of 250 in Arts faculty was found to be 120 lbs with standard deviation of 12 lbs. While the corresponding figure in the sample of 400 students from Science faculty were 124 lbs and 14 lbs. Is this difference significant?

Topic: Z proportion test (one sample and two samples)

Let's calculate the Z-score using the following formula:
Z = (X1 - X2) / sqrt((s1^2 / n1) + (s2^2 / n2))

Where: X1 = average weight of students in Arts faculty = 120 lbs X2 = average weight of students in Science faculty = 124 lbs s1 = standard deviation of Arts faculty = 12 lbs s2 = standard deviation of Science faculty = 14 lbs n1 = sample size of Arts faculty = 250 n2 = sample size of Science faculty = 400

Plugging in the values, we get:
Z = (120 - 124) / sqrt((12^2 / 250) + (14^2 / 400)) Z = -4 / 0.862 Z = -4.64
Looking up the Z-score in a standard normal distribution table, we find that the probability of getting a Z-score of -4.64 or lower is very low, approximately 0.000005.

8) Out of a sample of 100 residents in a certain area, 73 found their own homes. Test the hypothesis that the proportion of house owners is 80% against it is less than 80% at 5% level of significance.

To test the hypothesis that the proportion of house owners in the population is less than 80%, we can use a one-sample Z-test for proportions.

Let's define the null hypothesis H0 as the proportion of house owners in the population is 80%, and the alternative hypothesis Ha as the proportion of house owners in the population is less than 80%.

Let's calculate the Z-score using the following formula:

Z = (p - P0) / sqrt(P0*(1-P0) / n)
Where: p = proportion of house owners in the sample = 73/100 = 0.73 P0 = hypothesized proportion of house owners in the population = 0.80 n = sample size = 100

Plugging in the values, we get:
Z = (0.73 - 0.80) / sqrt(0.80*(1-0.80) / 100) Z = -0.39
Looking up the Z-score in a standard normal distribution table, we find that the probability of getting a Z-score of -0.39 or lower is approximately 0.35.

9) Random samples of 200 bolts manufactured by machine A and 100 bolts manufactured by machine B showed 19 and 5 defective bolts respectively. Is machine B better than machine A? Given α= 0.05.

To test if machine B is better than machine A, we can use a hypothesis test for the difference in proportions between two independent samples.
Let's define the null hypothesis H0 as the proportion of defective bolts produced by machine B is the same as or worse than that produced by machine A, and the alternative hypothesis Ha as the proportion of defective bolts produced by machine B is better than that produced by machine A.

Let's calculate the test statistic, which follows a standard normal distribution, using the following formula:

Z = (p1 - p2) / sqrt(p*(1-p)*(1/n1 + 1/n2))
Where: p1 = proportion of defective bolts in the sample from machine A = 19/200 = 0.095 p2 = proportion of defective bolts in the sample from machine B = 5/100 = 0.05 p = pooled proportion = (19+5) / (200+100) = 0.075 n1 = sample size from machine A = 200 n2 = sample size from machine B = 100

Plugging in the values, we get:
Z = (0.095 - 0.05) / sqrt(0.075*(1-0.075)*(1/200 + 1/100)) Z = 2.39

Using a standard normal distribution table, we find that the probability of getting a Z-score of 2.39 or higher is approximately 0.008.

10) In a village A out of a random sample of 1000 persons 100 were found to be a vegetarian. While in another village B out of 1500 persons 180 were found to be vegetarian. Do you find a significant difference in the food habits of the people of two villages? use a 5% level of significance.

To test the hypothesis, we will use the following formula for the test statistic:
z = (p1 - p2) / √(pooled proportion * (1 - pooled proportion) * (1/n1 + 1/n2))

where: pooled proportion = (x1 + x2) / (n1 + n2) x1 = number of vegetarians in village A = 100 n1 = sample size of village A = 1000 x2 = number of vegetarians in village B = 180 n2 = sample size of village B = 1500

Plugging in the values, we get:
pooled proportion = (100 + 180) / (1000 + 1500) = 0.1067
z = (0.1 - 0.12) / √(0.1067 * (1 - 0.1067) * (1/1000 + 1/1500)) z = -2.95
The calculated z-value of -2.95 is the test statistic.

At 5% level of significance for a two-tailed test, the critical value is ±1.96. Since the calculated z-value of -2.95 is less than the critical value of -1.96, we reject the null hypothesis.

Conclusion: There is sufficient evidence to conclude that there is a significant difference in the proportion of vegetarians between the two villages at 5% level of significance