**Topic: F-test:**

**1) Conduct 1-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**

Solution:

Step-1: -

First write the hypothesis statements as:

H\_0: No difference in variances.

H\_a: Difference in variances.

Step-2: -

Calculate the F-critical value. Here take the highest variance as the numerator and the lowest variance as the denominator:

F-Value= (Sigma1)2/(Sigma2)2

F-Value= 109.63/65.99

F-Value= 1.66

Step-3: -

Calculate the degrees of freedom as:

The degrees of freedom in the table will be the sample size -1, so for sample-1 it is 40 and for sample-2 it is 20.

Step-4: -

Choose the alpha level. As, alpha level was given in the question, so we may use the level of 0.025. This needs to be halved for the test, so use 0.0125.

Step-5: -

We will find the critical F-Value using the F-Table. We will use the table with 0.0125. Critical-F for (40,20) at alpha (0.0125) is 2.695.

Step-6: -

Compare the calculated value to the standard table value. If our calculated value is higher than the table value, then we may reject the null hypothesis. Here, 1.66 < 2.695. So, we cannot reject the null hypothesis.

**2) What is analysis of variance (ANOVA) and where it is used? Give the mathematical model of one-way ANOVA.**

Solution:

Analysis of Variance (ANOVA) is a statistical formula used to compare variances across the means (or average) of different groups. A range of scenarios use it to determine if there is any difference between the means of different groups.

Mathematical Model:

The mathematical model for one-way ANOVA can be represented as:

Yij = μ + τi + εij

Yij represents the jth observation in the ith group.

μ represents the overall population mean.

τi represents the effect of the ith group or treatment (deviation from the overall mean).

εij represents the random error term associated with each observation.

Hypothesis testing for small samples (t-distribution)

**3) Retail sales record shows that average monthly expenditure per family for a certain commodity was Rs. 255. A random sample of say 10 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?**

Solution:

Step 1:

Let's suppose the following null hypothesis (H0) and alternative hypothesis (Ha):

H0: The average monthly expenditure per family is Rs. 255.

Ha: The average monthly expenditure per family is different from Rs. 255.

Step 2:

Let's assume a significance level (α) of 0.05. This means we want to have a 95% confidence level for our hypothesis test.

Step 3:

we can use the t-test formula:

t = (mean - mu) / (sd / √n)

Where:

mean = sample mean (285)

mu = population mean under the null hypothesis (255)

sd = sample standard deviation (40)

n = sample size (10)

Plugging in the values:

t = (285 - 255) / (40 / √10) = 3.162

Step 4:

The degrees of freedom (df) = n-1 =10-1 = 9

Step 5:

We will find the critical t-Value using the t-Table. We will use the table with 0.025. Critical-t for df(9) is ±2.262.

The calculated test statistic (3.162) is greater than the positive critical value (2.262). Therefore, we reject the null hypothesis. These provides evidence to suggest that there is a significant change in the average monthly expenditure per family compared to the previously recorded value of Rs. 255.

**4) State the situation in which paired t-tests are used.**

Solution:

Paired t-tests are used when we want to compare the means of two related or dependent samples. In this type of situation, each observation in one sample is directly linked or paired with a corresponding observation in the other sample.

**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 75mm at a 5% level of significance and give a conclusion. The population standard deviation is known to be 9 mm.**

Solution:

State the hypotheses:

(H0): The mean diastolic blood pressure is 75 mm.

(Ha): The mean diastolic blood pressure is not equal to 75 mm.

Significance level:

The significance level (α) is given as 5% or 0.05.

Collect and summarize the data:

The sample size (n) is 81, and the sample mean is 79.2 mm. The population standard deviation (sd) is known to be 9 mm.

Conduct the hypothesis test:

Since the population standard deviation is known and the sample size is relatively large (n > 30), we can use the z-test.

The test statistic (z-score) can be calculated using the formula:

z = (mean - mu) / (sd / sqrt (n))

Putting in the values:

z = (79.2 - 75) / (9 / sqrt (81))

z = 4.2 / (9 / 9)

z = 4.2

Compare the p-value to the significance level. Since the p-value (p < 0.0001) is smaller than the significance level (a = 0.05), we reject the null hypothesis.

Based on the sample data, there is strong evidence to suggest that the mean diastolic blood pressure is significantly different from 75 mm.

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

Solution:

State the hypotheses:

(H0): The mean weight of the population is 2 kg.

(Ha): The mean weight of the population is not equal to 2 kg.

Significance level:

The significance level (α) is given as 5% or 0.05.

Collect and summarize the data:

The sample size (n) is 256, and the sample mean is 2.12 mm. The population standard deviation (sd) is 560gm (0.56).

Conduct the hypothesis test:

Since the population standard deviation is known and the sample size is relatively large (n > 30), we can use the z-test.

The test statistic (z-score) can be calculated using the formula:

z = (mean - mu) / (sd / sqrt (n))

Putting in the values:

z = (2.12 - 2) / (0.56 / sqrt (256))

z = 0.12 / (0.56 /16)

z = 0.12/0.035

z=3.43

Compare the p-value to the significance level (a). Since the p-value (p < 0.0001) is smaller than the significance level (a = 0.05), we reject the null hypothesis.

Based on the sample data, there is strong evidence to suggest that the mean weight of the population is significantly different from 2 kg.

**7) In a college there are two faculties Arts & Sciences. 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 Sciences faculty were 124 lbs. Is this different significance?**

Solution:

State the hypotheses:

(H0): The average weight of students in the Arts faculty is equal to the average weight of students in the Sciences faculty.

(Ha): The average weight of students in the Arts faculty is different from the average weight of students in the Sciences faculty.

Select the significance level:

The significance level (α) is not specified in the given problem. Let's assume a significance level of 5% or 0.05.

Collect and summarize the data:

For the Arts faculty: Sample size (n1) = 250, sample mean (x̄1) = 120 lbs, and sample standard deviation (s1) = 12 lbs.

For the Sciences faculty: Sample size (n2) = 400, sample mean (x̄2) = 124 lbs.

Conduct the hypothesis test:

Since we have two independent samples and the population standard deviations are unknown, we can use a two-sample t-test to compare the means.

The test statistic (t-score) can be calculated using the formula:

t = (x̄1 - x̄2) / sqrt((s1^2 / n1) + (s2^2 / n2))

Where:

x̄1 and x̄2 are the sample means for the Arts and Sciences faculties, respectively.

s1 and s2 are the sample standard deviations for the Arts and Sciences faculties, respectively.

n1 and n2 are the sample sizes for the Arts and Sciences faculties, respectively.

Plugging in the values:

t = (120 - 124) / sqrt((12^2 / 250) + (12^2 / 400))

t = -4 / sqrt(0.576 + 0.432)

t = -4 / sqrt(1.008)

t = -4 / 1.004

t = -3.98

Using the degrees of freedom (df) formula: df = (n1 - 1) + (n2 - 1) = 250 - 1 + 400 - 1 = 648, we can find the p-value associated with the t-score using a t-distribution table or a statistical calculator. The p-value is found to be very small, approximately p < 0.0001.

Based on the sample data, there is strong evidence to suggest that there is a significant difference in the average weights between the Arts and Sciences faculties. The average weight of students in one faculty is different from the average weight of students in the other faculty.

**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.**

Solutions:

State the hypotheses:

(H0): The proportion of house owners in the population is 80%.

(Ha): The proportion of house owners in the population is less than 80%.

Select the significance level:

The significance level (α) is given as 5% or 0.05.

Collect and summarize the data:

Out of a sample of 100 residents, 73 found their own homes. This gives us the sample proportion (p̂) = 73/100 = 0.73. The sample size (n) is 100.

Conduct the hypothesis test:

Calculate the test statistic, which is the z-score, using the formula:

z = (p̂ - p0) / sqrt((p0 \* (1 - p0)) / n)

Where:

p̂ is the sample proportion

p0 is the hypothesized population proportion under the null hypothesis

n is the sample size

Plugging in the values:

z = (0.73 - 0.80) / sqrt((0.80 \* (1 - 0.80)) / 100)

z = -0.07 / sqrt((0.80 \* 0.20) / 100)

z = -0.07 / sqrt(0.16 / 100)

z = -0.07 / sqrt(0.0016)

z = -0.07 / 0.04

z = -1.75

Calculate the p-value:

To find the p-value associated with the calculated z-score, we compare it to the standard normal distribution. Since the alternative hypothesis is one-sided (less than 80%), we find the area to the left of the z-score of -1.75 in the standard normal distribution using a table or a statistical calculator. The p-value is found to be approximately p = 0.0401.

Make a decision:

Compare the p-value to the significance level (α). Since the p-value (p = 0.0401) is smaller than the significance level (α = 0.05), we reject the null hypothesis.

Based on the sample data, there is sufficient evidence to suggest that the proportion of house owners in the population is less than 80% in the given area at a 5% level of significance.

**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 a=0.05**

Solutions:

State the hypotheses:

(H0): There is no difference in the proportion of defective bolts between machine A and machine B.

(Ha): Machine B has a lower proportion of defective bolts compared to machine A.

Select the significance level:

The significance level (α) is given as 0.05.

Collect and summarize the data:

For machine A: Sample size (n1) = 200, number of defective bolts (x1) = 19.

For machine B: Sample size (n2) = 100, number of defective bolts (x2) = 5.

Conduct the hypothesis test:

We will compare the proportions of defective bolts using a two-sample z-test for proportions. The test statistic is calculated as follows:

z = ((p̂1 - p̂2) - 0) / sqrt((p̂1 \* (1 - p̂1) / n1) + (p̂2 \* (1 - p̂2) / n2))

Where:

p̂1 is the sample proportion of defective bolts for machine A (x1 / n1)

p̂2 is the sample proportion of defective bolts for machine B (x2 / n2)

n1 and n2 are the sample sizes for machine A and machine B, respectively.

Plugging in the values:

p̂1 = 19/200 = 0.095

p̂2 = 5/100 = 0.05

n1 = 200

n2 = 100

z = ((0.095 - 0.05) - 0) / sqrt((0.095 \* (1 - 0.095) / 200) + (0.05 \* (1 - 0.05) / 100))

Calculate the p-value:

Calculate the p-value associated with the calculated z-score. Since the alternative hypothesis is one-sided (lower proportion for machine B), we find the area to the left of the z-score using a standard normal distribution table or a statistical calculator.

Make a decision:

Compare the p-value to the significance level (α). If the p-value is less than α, we reject the null hypothesis in favor of the alternative hypothesis. If the p-value is greater than α, we fail to reject the null hypothesis.

Based on the analysis, if the p-value is less than 0.05, we can conclude that there is sufficient evidence to suggest that machine B has a lower proportion of defective bolts compared to machine A. If the p-value is greater than 0.05, we do not have enough evidence to conclude that machine B is better than machine A in terms of producing defective bolts.

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

Solutions:

State the hypotheses:

(H0): There is no difference in the proportion of vegetarians between Village A and Village B.

(Ha): There is a significant difference in the proportion of vegetarians between Village A and Village B.

Select the significance level:

The significance level (α) is given as 5% or 0.05.

Collect and summarize the data:

For Village A: Sample size (n1) = 1000, number of vegetarians (x1) = 100.

For Village B: Sample size (n2) = 1500, number of vegetarians (x2) = 180.

Conduct the hypothesis test:

We will compare the proportions of vegetarians using a two-sample z-test for proportions. The test statistic is calculated as follows:

z = ((p̂1 - p̂2) - 0) / sqrt((p̂1 \* (1 - p̂1) / n1) + (p̂2 \* (1 - p̂2) / n2))

Where:

p̂1 is the sample proportion of vegetarians for Village A (x1 / n1)

p̂2 is the sample proportion of vegetarians for Village B (x2 / n2)

n1 and n2 are the sample sizes for Village A and Village B, respectively.

Plugging in the values:

p̂1 = 100/1000 = 0.1

p̂2 = 180/1500 = 0.12

n1 = 1000

n2 = 1500

z = ((0.1 - 0.12) - 0) / sqrt((0.1 \* (1 - 0.1) / 1000) + (0.12 \* (1 - 0.12) / 1500))

Calculate the p-value:

Calculate the p-value associated with the calculated z-score. Since the alternative hypothesis is two-sided (a significant difference), we find the area in the tails of the standard normal distribution that is greater than the absolute value of the z-score. This accounts for deviations in either direction.

Make a decision:

Compare the p-value to the significance level (α). If the p-value is less than α, we reject the null hypothesis in favor of the alternative hypothesis. If the p-value is greater than α, we fail to reject the null hypothesis.

Based on the analysis, if the p-value is less than 0.05, we can conclude that there is sufficient evidence to suggest a significant difference in the proportion of vegetarians between Village A and Village B. If the p-value is greater than 0.05, we do not have enough evidence to conclude a significant difference in food habits between the two villages.