Contents

Question set -1(STATISTICS) by(s.p)

1.Define F-statistics/distribution,

An F distribution is a probability distribution that results from comparing the variances of two samples or populations using the F statistic. It is the distribution of all possible F values for a specific combination of samples sizes that are being compared.

write down the pdf of F-statistic.

F-Distribution

Theorem 3.9.3 Let U and V be two independent chi-square random variables with n and d degrees of freedom, respectively. Define the random variable

$$F = \frac{U/n}{V/d}$$

The p.d.f. of F is

$$f(x) = \frac{\Gamma\left(\frac{n+d}{2}\right) n^{n/2} d^{d/2} x^{(n/2)-1}}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{d}{2}\right)(d+nx)^{(n+d)/2}}, \quad x > 0$$

or

The F distribution with `df1 =` ν_1 and `df2 =` ν_2 degrees of freedom has density

$$f(x) = \frac{\Gamma(\nu_1/2+\nu_2/2)}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} x^{\nu_1/2-1} \left(1 + \frac{\nu_1 x}{\nu_2}\right)^{-(\nu_1+\nu_2)/2}$$

or

Hence, the pdf of F is defined as

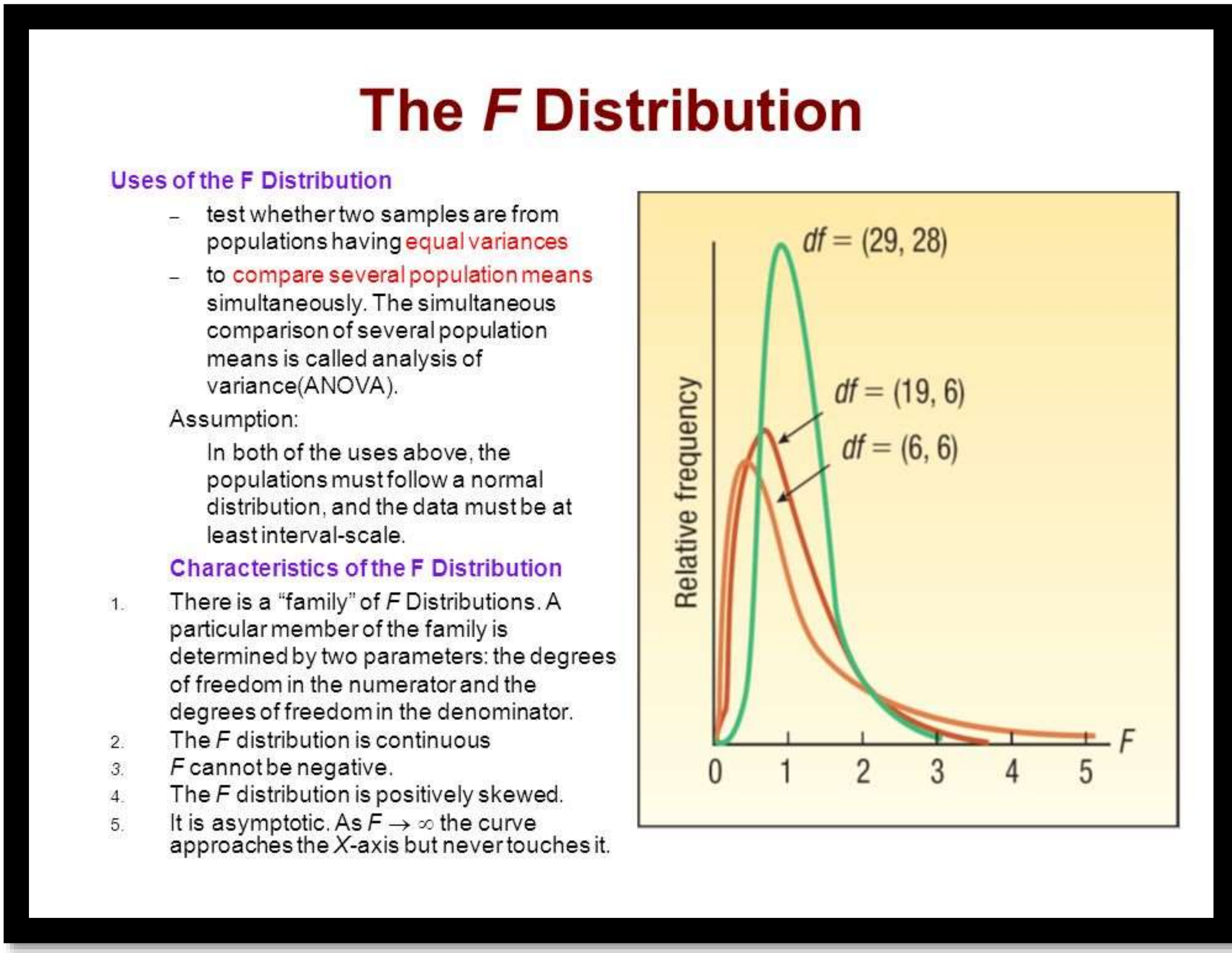
$$dP(F) = \frac{1}{B\left(\frac{v_1}{2},\frac{v_2}{2}\right)} \frac{\left(\frac{v_1}{v_2}F\right)^{\frac{v_1}{2}-1}}{\left(1+\frac{v_1}{v_2}F\right)^{\frac{v_1+v_2}{2}}} \, d\left(\frac{v_1}{v_2}F\right)$$

$$= \frac{1}{B\left(\frac{v_1}{2},\frac{v_2}{2}\right)} \frac{\left(\frac{v_1}{v_2}F\right)^{\frac{v_1}{2}-1}}{\left(1+\frac{v_1}{v_2}F\right)^{(v_1+v_2)/2}} \frac{v_1}{v_2} d(F)$$

$$\boxed{\Rightarrow f(F) = \frac{\left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}}}{B\left(\frac{v_1}{2},\frac{v_2}{2}\right)} \frac{(F)^{\frac{v_1}{2}-1}}{\left(1+\frac{v_1}{v_2}F\right)^{\frac{v_1+v_2}{2}}} \; ; \quad 0 \le F < \infty}$$

2. Mention the important applications/uses of F-distribution

- F-statistics has many applications in Statistics as well as any other scientific filed of research. However, some applications of F-statistic in Statistical theory are given below:
- 1. F-test for equality of Two Population variance
- 2. F-test for testing the significance of an observed multiple correlation coefficient
- 3. F-test for testing the significance of an observed sample correlation ratio
- 4. F-test for testing the linearity of Regression
- 5. F-test for testing equality of means
- 6. F is used for analysis of variance.



<u>what are the properties of F-distribution?</u>

The F-distribution has several important properties, including:

<u>**Skewness:**</u> The F-distribution is typically positively skewed, which means that it has a long right tail.

<u>**Symmetry**</u>**:** The F-distribution is symmetric only when the degrees of freedom for the numerator and denominator are equal.

<u>**Support: The**</u> F-distribution is defined over the non-negative real numbers.

Probability density function: The probability density function of the F-distribution is a continuous, smooth curve that approaches zero as x approaches negative infinity and as x approaches infinity.

Moments: The moments of the F-distribution are defined only when the degrees of freedom are greater than 2. The mean and variance of the F-distribution are given by:

Mean = df2 / (df2 - 2)

Variance = (2 * df2^2 * (df1 + df2 - 2)) / (df1 * (df2 - 2)^2 * (df2 - 4))

Relationship with other distributions: The F-distribution is a special case of the beta distribution, and it can be derived from the ratio of two independent chi-squared distributions.

Overall, the F-distribution is an important probability distribution in statistical analysis that has several useful properties for testing hypotheses and comparing populations.

3. Under usual notations prove that F= v2/2 [p –v2/2 - 1]

Example 16·20. *When $v_1 = 2$, show that the significance level of F corresponding to a significant probability p is :* $F = \frac{v_2}{2}\left[p^{-(2/v_2)} - 1\right]$ *where υ_1 an υ_2 have their usual meanings.*

Solution. When $v_1 = 2$, the p.d.f of $F(\upsilon_1\ \upsilon_2)$ distribution is:

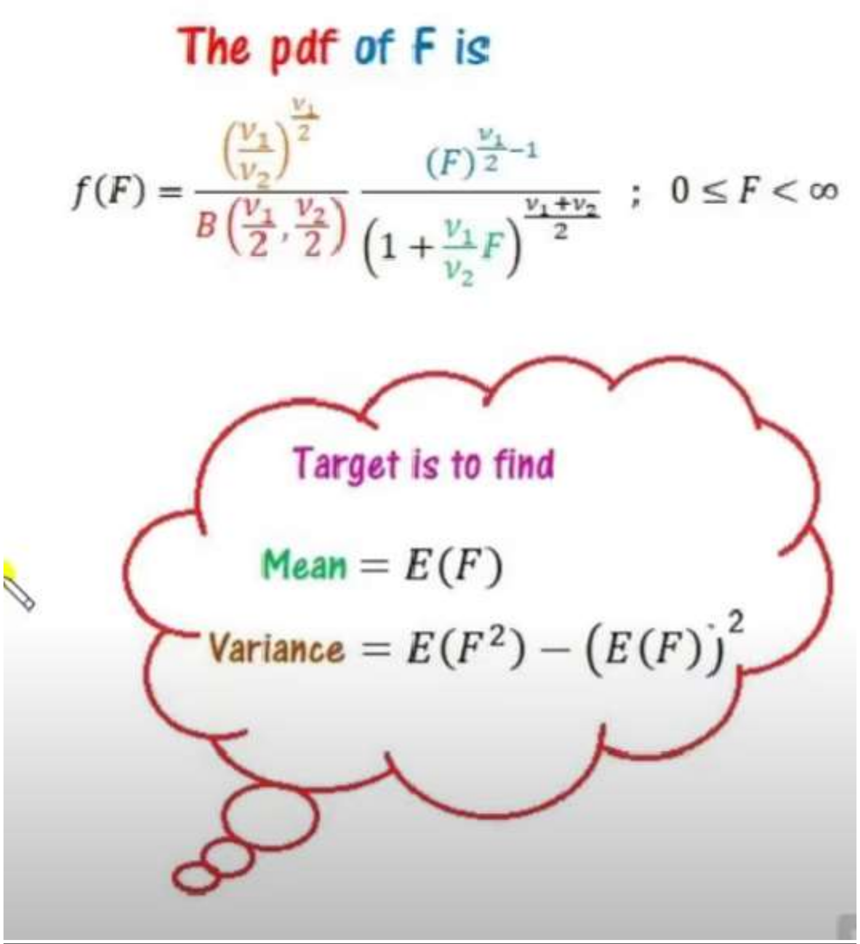
$$f(F) = \frac{1}{\left(1, \frac{v_2}{2}\right)} \cdot \frac{2}{v_2} \cdot \frac{1}{\left(1 + \frac{2}{v_2}F\right)^{(v_2/2)+1}} \ ; 0 < F < \infty$$

$$= \frac{\Gamma\left(\frac{v_2}{2}+1\right)}{\Gamma(1)\Gamma(v_2/2)} \times \frac{2/v_2}{\left(\frac{2}{v_2}\right)^{(v_2/2)+1}\left(F + \frac{v_2}{2}\right)^{(v_2/2)+1}} = \frac{\left(\frac{v_2}{2}\right)^{(v_2/2)+1}}{\left(F + \frac{v_2}{2}\right)^{(v_2/2)+1}}$$

Hence, $$p = \int_F^\infty f(F)dF = \left[\left(\frac{v_2}{2}\right)\right]^{(v_2/2)+1} \times \int_F^\infty \frac{dF}{\left(F + \frac{v_2}{2}\right)^{(v_2/2)+1}}$$

$$= \left(\frac{v_2}{2}\right)^{(v_2/2)+1} \times \left|\frac{\left(F + \frac{v_2}{2}\right)^{-(v_2/2)}}{-\frac{v_2}{2}}\right|_F^\infty = \left[\frac{\left(\frac{v_2}{2}\right)}{F + \frac{v_2}{2}}\right]^{(v_2/2)} = \frac{1}{\left(1 + \frac{2}{v_2}F\right)^{(v_2/2)}}$$

4.Find mean, median, variance of F-distribution.?



$$E(F^r) = \int_0^\infty F^r f(F)\,dF$$

$$= \int_0^\infty F^r \frac{\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}}{B\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)} \frac{(F)^{\frac{\nu_1}{2}-1}}{\left(1+\frac{\nu_1}{\nu_2}F\right)^{\frac{\nu_1+\nu_2}{2}}}\,dF$$

$$= \frac{\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}}{B\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)} \int_0^\infty \frac{F^{r+\frac{\nu_1}{2}-1}}{\left(1+\frac{\nu_1}{\nu_2}F\right)^{\frac{\nu_1}{2}+\frac{\nu_2}{2}}}\,dF$$

$$\boxed{\text{Take } \frac{\nu_1}{\nu_2}F = y \;\Rightarrow\; dF = \frac{\nu_2}{\nu_1}dy}$$

$$= \frac{\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}}{B\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)} \int_0^\infty \frac{\left(\frac{\nu_2}{\nu_1}y\right)^{r+\frac{\nu_1}{2}-1}}{(1+y)^{\frac{\nu_1}{2}+\frac{\nu_2}{2}}}\left(\frac{\nu_2}{\nu_1}dy\right)$$

$$= \frac{\left(\frac{\nu_2}{\nu_1}\right)^{r}}{B\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)} \int_0^\infty \frac{(y)^{r+\frac{\nu_1}{2}-1}}{(1+y)^{\frac{\nu_1}{2}+\frac{\nu_2}{2}}}\,dy$$

B(m,n)

m>0

n>0

$$= \frac{\left(\frac{\nu_2}{\nu_1}\right)^{r}}{B\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)} \int_0^\infty \frac{(y)^{r+\frac{\nu_1}{2}-1}}{(1+y)^{\frac{\nu_1}{2}+r+\frac{\nu_2}{2}-r}}\,dy$$

$$= \frac{\left(\frac{\nu_2}{\nu_1}\right)^{r}}{B\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)} B\left(\frac{\nu_1}{2}+r,\frac{\nu_2}{2}-r\right) \quad ; \quad \nu_2 > 2r$$

$$= \frac{\left(\frac{\nu_2}{\nu_1}\right)^{r}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)/\cancel{\Gamma\left(\frac{\nu_1}{2}+\frac{\nu_2}{2}\right)}} \frac{\Gamma\left(\frac{\nu_1}{2}+r\right)\Gamma\left(\frac{\nu_2}{2}-r\right)}{\cancel{\Gamma\left(\frac{\nu_1}{2}+\frac{\nu_2}{2}\right)}}$$

$$= \left(\frac{\nu_2}{\nu_1}\right)^{r} \frac{\Gamma\left(\frac{\nu_1}{2}+r\right)\Gamma\left(\frac{\nu_2}{2}-r\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \quad ; \quad \nu_2 > 2r$$

$$E(F^r) = \left(\frac{\nu_2}{\nu_1}\right)^{r} \frac{\Gamma\left(\frac{\nu_1}{2}+r\right)\Gamma\left(\frac{\nu_2}{2}-r\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \quad ; \quad \nu_2 > 2r$$

$$E(F) = \left(\frac{\nu_2}{\nu_1}\right) \frac{\Gamma\left(\frac{\nu_1}{2}+1\right)\Gamma\left(\frac{\nu_2}{2}-1\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}$$

$$= \left(\frac{\nu_2}{\nu_1}\right) \frac{\left(\frac{\nu_1}{2}\right)\cancel{\Gamma\left(\frac{\nu_1}{2}\right)}\cancel{\Gamma\left(\frac{\nu_2}{2}-1\right)}}{\cancel{\Gamma\left(\frac{\nu_1}{2}\right)}\left(\frac{\nu_2}{2}-1\right)\cancel{\Gamma\left(\frac{\nu_2}{2}-1\right)}}$$

$$= \frac{\nu_2}{\nu_1}\frac{\left(\frac{\nu_1}{2}\right)}{\left(\frac{\nu_2}{2}-1\right)}$$

$$= \frac{\nu_2}{\nu_2-2} \quad ; \quad \nu_2 > 2$$

$$E(F^2) = \left(\frac{\nu_2}{\nu_1}\right)^2 \frac{\Gamma\left(\frac{\nu_1}{2}+2\right)\Gamma\left(\frac{\nu_2}{2}-2\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \quad ; \quad \nu_2 > 4$$

$$= \left(\frac{\nu_2}{\nu_1}\right)^2 \frac{\left(\frac{\nu_1}{2}+1\right)\left(\frac{\nu_1}{2}\right)\cancel{\Gamma\left(\frac{\nu_1}{2}\right)}\cancel{\Gamma\left(\frac{\nu_2}{2}-2\right)}}{\cancel{\Gamma\left(\frac{\nu_1}{2}\right)}\left(\frac{\nu_2}{2}-1\right)\left(\frac{\nu_2}{2}-2\right)\cancel{\Gamma\left(\frac{\nu_2}{2}-2\right)}}$$

$$= \left(\frac{\nu_2}{\nu_1}\right)^2 \frac{\left(\frac{\nu_1}{2}+1\right)\left(\frac{\nu_1}{2}\right)}{\left(\frac{\nu_2}{2}-1\right)\left(\frac{\nu_2}{2}-2\right)}$$

$$= \frac{\nu_2^2(\nu_1+2)}{\nu_1(\nu_2-2)(\nu_2-4)} \quad ; \quad \nu_2 > 4$$

Hence,

$$\text{Mean} = \frac{v_2}{v_2-2} \quad ; \quad v_2 > 2$$

$$\begin{aligned}
\text{Variance} &= E(F^2) - \left(E(F)\right)^2 \\
&= \frac{v_2^2(v_1+2)}{v_1(v_2-2)(v_2-4)} - \left(\frac{v_2}{v_2-2}\right)^2 \\
&= \frac{v_2^2}{v_2-2}\left[\frac{v_1+2}{v_1(v_2-4)} - \frac{1}{v_2-2}\right] \\
&= \frac{2v_2^2(v_1+v_2-2)}{v_1(v_2-2)^2(v_2-4)} \quad ; \quad v_2 > 4
\end{aligned}$$

5.Establish of relation between T and F-distribution ?

Applied Statistics by the same authors.]

16·8. RELATION BETWEEN t AND F DISTRIBUTIONS

In F-distribution with (υ_1, υ_2) *d.f.* [*c.f.* 16·13 (*a*)], take $\nu_1 = 1$, $\nu_2 = \nu$ and $t^2 = F$, *i.e.*, $dF = 2t\,dt$. Thus, the probability differential of F transforms to :

$$dG(t) = \frac{(1/\nu)^{1/2}}{B\left(\frac{1}{2}, \frac{\nu}{2}\right)} \cdot \frac{\left(t^2\right)^{(1/2)-1}}{\left(1+\frac{t^2}{\nu}\right)^{(\nu+1)/2}}\, 2t\,dt,\ 0 \le t^2 < \infty$$

$$= \frac{1}{\sqrt{\nu}\, B\left(\frac{1}{2}, \frac{\nu}{2}\right)} \cdot \frac{1}{\left(1+\frac{t^2}{\nu}\right)^{(\nu+1)/2}}\, dt,\ -\infty < t < \infty$$

the factor 2 disappearing since the total probability in the range $(-\infty, \infty)$ is unity. This is the probability function of Student's t-distribution with ν *d.f.* Hence, we have the following relation between t and F distributions.

If a statistic t follows Student's t distribution with n d.f., then t^2 follows Snedecor's F-distribution with (1, n) d.f. Symbolically,

$$\left.\begin{array}{ll} \text{if} & t \sim t_{(n)} \\ \text{then} & t^2 \sim F_{(1,n)} \end{array}\right\} \qquad \ldots (16{\cdot}21)$$

6. Establish of relation between X2 and F-distribution?

16·9. RELATION BETWEEN F AND χ^2 DISTRIBUTION

In $F(n_1, n_2)$ distribution if we let $n_2 \to \infty$, then $\chi^2 = n_1 F$ follows χ^2-distribution with n_1 *d.f.*

Proof. $$f(F) = \frac{(n_1/n_2)^{n_1/2} F^{(n_1/2)-1}}{\Gamma(n_1/2)\Gamma(n_2/2)} \cdot \frac{\Gamma[(n_1+n_2)/2]}{\left(1+\frac{n_1}{n_2}F\right)^{(n_1+n_2)/2}}, \quad 0 < F < \infty \qquad \ldots(16\cdot22)$$

In the limit as $n_2 \to \infty$, we have

$$\frac{\Gamma[(n_1+n_2)/2]}{n_2^{n_1/2}\Gamma(n_2/2)} \to \frac{(n_2/2)^{n_1/2}}{n_2^{n_1/2}} = \frac{1}{2^{n_1/2}} \qquad \ldots(*)$$

$$\left[\because \frac{\Gamma(n+k)}{\Gamma(n)} \to n^k \text{ as } n \to \infty. (c.f. \text{ Remark below})\right]$$

Also $$\lim_{n_2\to\infty}\left(1+\frac{n_1}{n_2}F\right)^{(n_1+n_2)/2} = \lim_{n_2\to\infty}\left[\left(1+\frac{n_1}{n_2}F\right)^{n_2}\right]^{1/2} \times \lim_{n_2\to\infty}\left(1+\frac{n_1}{n_2}F\right)^{n_1/2}$$

$$= \exp(n_1F/2) = \exp(\chi^2/2) \qquad (\because n_1F = \chi^2) \qquad \ldots(**)$$

Hence, in the limit, on using (*) and (**), the *p.d.f.* of $\chi^2 = n_1F$ becomes :

$$dP(\chi^2) = \frac{(n_1/2)^{n_1/2} e^{-\chi^2/2}}{\Gamma(n_1/2)} \left(\frac{\chi^2}{n_1}\right)^{(n_1/2)-1} d\left(\frac{\chi^2}{n_1}\right)$$

$$= \frac{1}{2^{n_1/2}\Gamma(n_1/2)} \cdot e^{-\chi^2/2} (\chi^2)^{(n_1/2)-1} d\chi^2, \; 0 < \chi^2 < \infty$$

which is the p.d.f. of chi-square distribution with n_1 *d.f.*

Remark. $$\lim_{n\to\infty}\frac{\Gamma(n+k)}{\Gamma(n)} = \lim_{n\to\infty}\frac{(n+k-1)!}{(n-1)!}, \text{ (for large } n) = \lim_{n\to\infty}\frac{\sqrt{2\pi}\, e^{-(n+k-1)}(n+k-1)^{n+k-(1/2)}}{\sqrt{2\pi}\, e^{-(n-1)}(n-1)^{n-(1/2)}}$$

(On using Stirling's approximation for $n!$ as $n \to \infty$.)

$$= e^{-k}\lim_{n\to\infty}\frac{n^{n+k-\frac{1}{2}}\left(1+\frac{k-1}{n}\right)^{n+k-\frac{1}{2}}}{n^{n-\frac{1}{2}}\left(1-\frac{1}{n}\right)^{n-\frac{1}{2}}} = e^{-k}n^k \frac{\lim_{n\to\infty}\left(1+\frac{k-1}{n}\right)^n \lim_{n\to\infty}\left(1+\frac{k-1}{n}\right)^{k-\frac{1}{2}}}{\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{-\frac{1}{2}}}$$

Question set -2(STATISTICS) by(s.p)

describe the test procedure to test the folloing hypothesis

1.H1 : μ =0 vs H2 : μ ≠ 0

To test the hypothesis H1: μ = 0 against the alternative hypothesis H2: μ ≠ 0, where μ represents the population mean, we can perform a two-tailed hypothesis test. The test procedure involves the following steps:

Step 1: State the null and alternative hypotheses:

Null hypothesis (H1): The population mean is equal to 0, $\mu = 0$.

Alternative hypothesis (H2): The population mean is not equal to 0, $\mu \neq 0$.

Step 2: Set the significance level (α):

Choose a significance level (α) to determine the probability of making a Type I error. Common choices for α are 0.05 or 0.01, depending on the desired level of confidence.

Step 3: Collect and summarize the data:

Obtain a random sample from the population of interest and calculate the sample mean ($\bar{x}$) and sample standard deviation (s) from the data.

Step 4: Calculate the test statistic:

The test statistic for a two-tailed test of the population mean is the z-score. The formula for the z-score is given by:

$z = (\bar{x} - \mu_0) / (s / \sqrt{n})$

Where:

$\bar{x}$ = Sample mean

μ_0 = Assumed population mean under the null hypothesis (here, $\mu_0 = 0$)

s = Sample standard deviation

n = Sample size

Step 5: Find the critical value(s):

For a two-tailed test, the critical values will be based on the chosen significance level (α) and the appropriate probability distribution. The z-critical values for a two-tailed test at a 95% confidence level ($\alpha = 0.05$) are approximately ±1.96.

Step 6: Make a decision:

Compare the calculated z-score from Step 4 with the critical value(s) from Step 5. If the calculated z-score falls within the critical value range, fail to reject the null hypothesis (H1). If the calculated z-score falls outside the critical value range, reject the null hypothesis (H1) in favor of the alternative hypothesis (H2).

Step 7: Draw the conclusion:

Based on the decision made in Step 6, draw a conclusion about the population mean μ.

If the null hypothesis (H1) is not rejected, we do not have sufficient evidence to claim that the population mean is significantly different from 0.

If the null hypothesis (H1) is rejected, we have sufficient evidence to suggest that the population mean is significantly different from 0.

Remember that the conclusion is based on the sample data and is subject to the chosen significance level and the assumptions of the test.

2)Ho: $\sigma1^2 = \sigma2^2$,

To test the hypothesis Ho: $\sigma1^2 = \sigma2^2$, where $\sigma1^2$ represents the variance of population 1 and $\sigma2^2$ represents the variance of population 2, we can perform an F-test for equality of variances. This test helps determine whether the variances of two populations are significantly different or not.

The test procedure involves the following steps:

Step 1: State the null and alternative hypotheses:

Null hypothesis (Ho): The variance of population 1 ($\sigma1^2$) is equal to the variance of population 2 ($\sigma2^2$), i.e., $\sigma1^2 = \sigma2^2$. Alternative hypothesis: The variance of population 1 ($\sigma1^2$) is not equal to the variance of population 2 ($\sigma2^2$), i.e., $\sigma1^2 \neq \sigma2^2$.

Step 2: Set the significance level (α):

Choose a significance level (α) to determine the probability of making a Type I error. Common choices for α are 0.05 or 0.01, depending on the desired level of confidence.

Step 3: Collect and summarize the data:

Obtain random samples from the two populations of interest and calculate the sample variances ($s1^2$ and $s2^2$) from the data.

Step 4: Calculate the test statistic:

The test statistic for the F-test is the ratio of the larger sample variance to the smaller sample variance. The formula for the F-test statistic is given by:

$F = s1^2 / s2^2$ (if $s1^2 > s2^2$) or $F = s2^2 / s1^2$ (if $s2^2 > s1^2$)

Step 5: Find the critical value:

The critical value for the F-test is based on the chosen significance level (α), the degrees of freedom for the two samples, and the probability distribution. You can find the critical value from an F-distribution table.

Step 6: Make a decision:

Compare the calculated F-test statistic from Step 4 with the critical value from Step 5. If the calculated F-statistic falls within the critical value range, fail to reject the null hypothesis (Ho). If the calculated F-statistic falls outside the critical value range, reject the null hypothesis (Ho) in favor of the alternative hypothesis.

Step 7: Draw the conclusion:

Based on the decision made in Step 6, draw a conclusion about the equality of variances between the two populations.

- If the null hypothesis (Ho) is not rejected, we do not have sufficient evidence to claim that the variances of the two populations are significantly different.
- If the null hypothesis (Ho) is rejected, we have sufficient evidence to suggest that the variances of the two populations are significantly different.

Remember that the conclusion is based on the sample data and is subject to the chosen significance level and the assumptions of the test.

3) Describe how you will test the following hypothesis Ho: p= p2………………. Pk (k>2)

To test the hypothesis Ho: p = p2 = ... = Pk (k > 2), where p, p2, ..., Pk are population proportions, you can use a **chi-square goodness-of-fit test**. This test compares the observed frequencies of a categorical variable to the expected frequencies under the null hypothesis.

Here's a step-by-step guide on how to perform the chi-square goodness-of-fit test for this hypothesis:

Step 1: State the null and alternative hypotheses:

Null hypothesis (Ho): p = p2 = ... = Pk (All population proportions are equal).

Alternative hypothesis (Ha): At least one population proportion is different from the others.

Step 2: Collect and organize the data:

You need observed frequencies for each category and the total sample size (n). The data should be categorical, divided into k categories.

Step 3: Calculate the expected frequencies:

Under the null hypothesis, where all population proportions are equal, the expected frequency for each category is given by the formula:

Expected Frequency (E) = (Total Sample Size * Proportion assumed under Ho for that category)

In this case, since all proportions are assumed to be equal, you can set E = n / k for each category.

Step 4: Calculate the chi-square test statistic:

Chi-Square = Σ [(Observed Frequency - Expected Frequency)^2 / Expected Frequency]

Sum the values of [(Observed Frequency - Expected Frequency)^2 / Expected Frequency] for all k categories.

Step 5: Determine the degrees of freedom (df):

df = k - 1 (where k is the number of categories).

Step 6: Find the critical value:

At a given significance level (e.g., 0.05), find the critical value from the chi-square distribution table with df degrees of freedom.

Step 7: Compare the test statistic with the critical value:

If the test statistic is greater than the critical value, reject the null hypothesis. If it is smaller, fail to reject the null hypothesis.

Step 8: Draw the conclusion:

Based on the comparison, draw a conclusion about the null hypothesis. If the null hypothesis is rejected, it suggests that at least one population proportion is different from the others.

Note: The chi-square goodness-of-fit test assumes that the expected frequency for each category is at least 5. If this assumption is not met, you might need to consider combining categories or using an alternative test.