F- distribution (Test for variance)

(Snedecor's F- distribution)

Suppose that two independent normal populations are of interest, when the population means and variances are unknown. We wish to test hypothesis about the equality of the two variances, say $H_0 : \sigma_1^2 = \sigma_2^2$, and $H_1 : \sigma_1^2 \neq \sigma_2^2$

We use

$$F = \frac{S_1^2}{S_2^2}$$

where S_1 & S_2 are sample SD. and Nr > Dr.

Applications :-

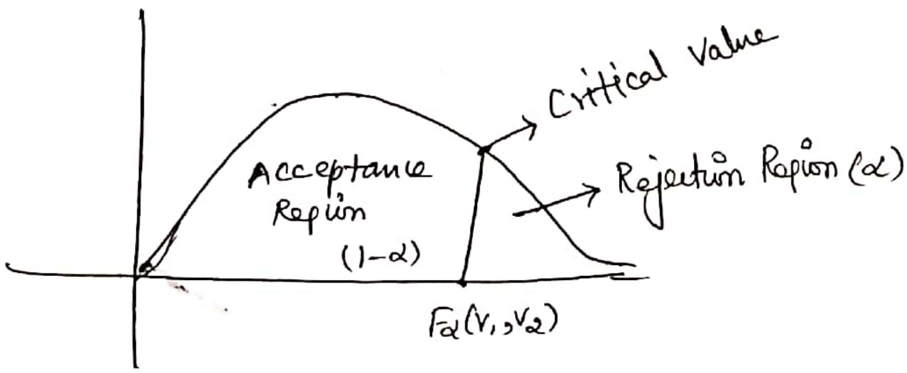
F- test is used to test,

1) whether two independent samples have been drawn from the normal populations with same variance σ^2.

2) whether the two independent estimates of the pop. Variance are homogeneous or not.

PROPERTIES

1) The probability curve of the F-distribution is roughly sketched in the fig.



2) The range of F distribution is 1 to ∞.

3) The value of F is always greater than 1.

Working Procedure :

* Set up the null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$

* $H_1 : \sigma_1^2 \neq \sigma_2^2$

* LOS = α

* DF

* Test statistics $F = \frac{S_1^2}{S_2^2}$; $\frac{\text{Greater variance}}{\text{Smaller variance}}$

→ unbiased estimate of population variance

where $S_1^2 = \frac{\Sigma(x-\bar{x})^2}{n_1 - 1}$; $S_2^2 = \frac{\Sigma(y-\bar{y})^2}{n_2-1}$

$$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1} \qquad S_2^2 = \frac{n_2 s_2^2}{n_2 - 1}$$

* Conclusion.

Problems under F-test :

1. A sample of size 13 gave an estimated population variance of 3.0, while another sample of size 15, gave an estimate of 2.5, Could both samples be from population with the same variance ?

Soln:

Given $n_1 = 13$ $n_2 = 15$

$\sigma_1^2 = 3.0$ $\sigma_2^2 = 2.5$

1. $H_0 : \sigma_1^2 = \sigma_2^2$

2. $H_1 : \sigma_1^2 \neq \sigma_2^2$

3. LOS = 5%

4. D.F

$\nu_1 = n_1 - 1 = 13 - 1 = 12$

$\nu_2 = n_2 - 1 = 15 - 1 = 14$

5. Test Statistics $F = \frac{S_1^2}{S_2^2}$ or $\frac{\sigma_1^2}{\sigma_2^2}$

$= \frac{3}{2.5} = 1.2 > 1$

6. Table value $F(\nu_1, \nu_2) = F(12, 14) = 2.53$

7. Conclusion :

Cal F < Tab F

$\therefore$ We can accept Ho.

2. In one sample of 10 observations, the sum of the squares of the deviations of the sample values from the sample mean was 120 and in another sample of 12 observations it was 314. Test whether this difference is significant at 5% LOS.

Solution :

$n_1 = 10$, $n_2 = 12$

$\Sigma(x-\bar{x})^2 = 120$; $\Sigma(y-\bar{y})^2 = 314$

1. $H_0 : \sigma_1^2 = \sigma_2^2$

2. $H_1 : \sigma_1^2 \neq \sigma_2^2$ (Two tailed)

3. LOS = 5%

4. DF = $n_1 - 1 = 10 - 1 = 9 = \nu_1$

$n_2 - 1 = 12 - 1 = 11 = \nu_2$

5. Test statistics $F = \frac{S_1^2}{S_2^2}$

$S_1^2 = \frac{\Sigma(x-\bar{x})^2}{n_1 - 1} = \frac{120}{9} = 13.33$

$$S_2^2 = \frac{\Sigma(y-\bar{y})^2}{n_2-1} = \frac{314}{11} = 28.5$$

Here $S_2^2 > S_1^2$

$$\therefore \quad F = \frac{S_2^2}{S_1^2} = \frac{28.5}{13.33} = 2.13$$

6. Table value $F(\nu_2, \nu_1) = \cancel{F(9, 11)} = F(11, 9) = 3.07$

Conclusion : Cal. F < Tab. F

Hence we accept Ho

3. Two random samples of 11 and 9 items shows the sample SD of their weights as 0.8 and 0.5 respy. Assuming that the weight distribution are normal. Test the hypothesis that the two variances are equal or not.

Soln:

$n_1 = 11, \quad n_2 = 9$

$s_1^2 = 0.8 \quad s_2^2 = 0.5$

1. $H_0 : \sigma_1^2 = \sigma_2^2$

2. $H_1 : \sigma_1^2 \neq \sigma_2^2$

3. LOS = 5 %

4. DF = $\nu_1 = n_1 - 1 = 11 - 1 = 10$

   $\nu_2 = n_2 - 1 = 9 - 1 = 8$

5. Test Statistics $F = \frac{S_1^2}{S_2^2}$ where

$$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1} = \frac{11(0.8)^2}{10} = 0.704$$

$$S_2^2 = \frac{n_2 s_2^2}{n_2 - 1} = \frac{9(0.5)^2}{8} = 0.28$$

Here $S_1^2 > S_2^2$

$$\therefore F = \frac{S_1^2}{S_2^2} = \frac{0.704}{0.28} = 2.51$$

6. Table value of $F(\gamma_1, \gamma_2) = F(10, 8) = 3.35$

7. Conclusion :

Cal F < tab F

So we can accept Ho.

(4) A group of 10 rats fed on diet A and another group of 8 rats fead on diet B, recorded the following increase in weights.

Diet A :	5	6	8	1	12	4	3	9	6	10
Diet B :	2	3	6	8	10	1	2	8	-	-

Find if the variances are significantly difference.

Soln:

Gn $n_1 = 10$, $n_2 = 8$

1. $H_0 : \sigma_1^2 = \sigma_2^2$
2. $H_1 : \sigma_1^2 \neq \sigma_2^2$ (Two tailed test)
3. LOS = 5 %
4. DF : $\gamma_1 = n_1 - 1 = 10 - 1 = 9$
   $\gamma_2 = n_2 - 1 = 8 - 1 = 7$
5. Test Statistics $F = \frac{S_1^2}{S_2^2}$

where $S_1^2 = \frac{\sum(x-\bar{x})^2}{n_1 - 1}$; $S_2^2 = \frac{\sum(y-\bar{y})^2}{n_2 - 1}$

x	5	6	8	1	12	4	3	9	6	10
$(x-\bar{x})^2$	1.96	0.16	2.56	29.16	31.36	5.76	11.56	6.76	0.16	12.96

$$\bar{x} = \frac{\Sigma x}{n} = 6.4$$

$$\Sigma(x-\bar{x})^2 = 102.34$$

y	2	3	6	8	10	1	2	8
$(y-\bar{y})^2$	9	4	1	9	25	16	9	9

$$\bar{y} = \frac{\Sigma y}{n} = 5$$

$$\Sigma(y-\bar{y})^2 = 82$$

$$S_1^2 = \frac{\Sigma(x-\bar{x})^2}{n_1-1} = \frac{102.34}{9} = 11.37$$

$$S_2^2 = \frac{\Sigma(y-\bar{y})^2}{n_2-1} = \frac{82}{7} = 11.71$$

Here $S_2^2 > S_1^2$

$$F = \frac{S_2^2}{S_1^2} = \frac{11.71}{11.37} = 1.03 > 1$$

6. Table value of $F(\gamma_1, \gamma_2) = F(9,9) = 3.29$

7. Conclusion :

Cal F < tab F

So we can accept H_0.

5. Two independent samples of sizes of 9 and 7 from a normal population had the following values of the variables.

Sample I	18	13	12	15	12	14	16	14	15
Sample II	16	19	13	16	18	13	15		

Do the estimates of the population variance differ significantly at 5% Level?

Soln:

$n_1 = 9, \quad n_2 = 7$

$$\bar{x} = \frac{18+13+12+15+12+14+16+14+15}{9} = 14.33$$

$$\bar{y} = \frac{16+19+13+16+18+13+15}{7} = 15.71$$

x	18	13	12	15	12	14	16	14	15
$(x-\bar{x})^2$	13.469	1.769	5.429	0.449	5.429	0.109	2.789	0.109	0.449

$\Sigma(x-\bar{x})^2 = 30.001$

y	16	19	13	16	18	13	15
$(y-\bar{y})^2$	0.084	10.824	7.344	0.084	5.244	7.344	0.504

$\Sigma(y-\bar{y})^2 = 31.428$

$$S_1^2 = \frac{\Sigma(x-\bar{x})^2}{n_1-1} = \frac{30.001}{8} = 3.75$$

$$S_2^2 = \frac{\Sigma(y-\bar{y})^2}{n_2-1} = \frac{31.428}{6} = 5.238$$

Scanned by CamScanner

1. $H_0 : \sigma_1^2 = \sigma_2^2$

2. $H_1 : \sigma_1^2 \neq \sigma_2^2$

3. LOS = 5 %

4. DF = $\gamma_1 = n_1 - 1 = 8$

$\gamma_2 = n_2 - 1 = 6.$

5. Test statistics $F = \frac{S_2^2}{S_1^2} = \frac{5.238}{3.75} = 1.3968$

6. Table value of $F(6,8) = 3.58$

7. Conclusion : Cal F < Tab F

So we can accept H_0.

6. Two random Sample of sizes 8 and 7 had the following values of the variables

Sample A	9	11	13	11	15	9	12	14
Sample B	10	12	10	14	9	8	10	

Do the estimates of the population variance differ significantly?

Soln: Gn $n_1 = 8$ $n_2 = 7$

$$\bar{x} = \frac{9+11+13+11+15+9+12+14}{8} = 11.75$$

$$\bar{y} = \frac{10+12+10+14+9+8+10}{7} = 10.42$$

x	9	11	13	11	15	9	12	14
$(x-\bar{x})^2$	7.562	0.562	1.562	0.562	10.562	7.562	0.0625	5.062

Scanned by CamScanner

$\Sigma(x-\bar{x})^2 = 33.496$

y	10	12	10	14	9	8	10
$(y-\bar{y})^2$	0.176	2.496	0.176	12.816	2.016	5.856	0.176

$\Sigma(y-\bar{y})^2 = 23.712$

$$S_1^2 = \frac{\Sigma(x-\bar{x})^2}{n_1-1} = \frac{33.496}{7} = 4.785$$

$$S_2^2 = \frac{\Sigma(y-\bar{y})^2}{n_2-1} = \frac{23.712}{6} = 3.952$$

1. $H_0 : \sigma_1^2 = \sigma_2^2$

2. $H_1 : \sigma_1^2 \neq \sigma_2^2$

3. LOS = 5%

4. DF $= \gamma_1 = n_1 - 1 = 7$

   $\gamma_2 = n_2 - 1 = 6$

5. $F = \frac{S_1^2}{S_2^2} = \frac{4.785}{3.952} = 1.21$

6. Table value $F(7,6) = 4.21$

7. Conclusion cal $F <$ Tab F

   $\therefore$ we can accept H_0.