Snedecor's F-test of Significance
(Small Samples)

Two samples came from the same population or from population with same variances.

If s_1^2 and s_2^2 are the variances of two samples of sizes n_1 and n_2 respectively. The estimate of the population variances of these two samples are respectively $S_1^2 = \frac{n_1 s_1^2}{n_1 - 1}$ and $S_2^2 = \frac{n_2 s_2^2}{n_2 - 1}$. $\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$ d.f. of freedom.

We want to test if these estimates S_1^2 and S_2^2 are significantly different or if the samples may be regarded as drawn from the same population or from two population with same Variance.

Let $F = \dfrac{S_1^2}{S_2^2} = \dfrac{\dfrac{n_1 s_1^2}{n_1 - 1}}{\dfrac{n_2 s_2^2}{n_2 - 1}}$

s_1^2, s_2^2 are Sample Variances

Then the F- distribution:

$$y = f(F) = \frac{k\, F^{\frac{\nu_1 - 2}{2}}}{(\nu_1 F + \nu_2)^{\frac{\nu_1 + \nu_2}{2}}}$$

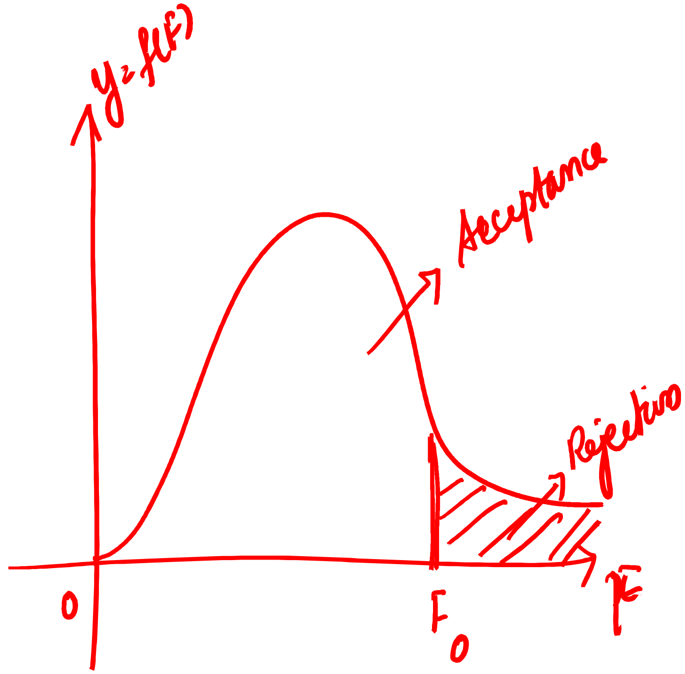
where ν_1 and ν_2 are the d.f of two estimates &

k can be got from $\int_0^{\infty} f(F)\,dF = 1$.

If $S_1^2 = S_2^2$ then $F = 1$

Hence how far any observed value of F differs from unity, Consistent with our assumption of equality of the population Variances.

$F_0 \rightarrow$ Critical Value



Area under the curve $y = f(F)$.

to the right of F_0.

So if $F_0 > F$ then we accept the null hypothesis

$F_0 < F$ or $F > F_0$ then Reject the null hypothesis

In setting the greater value for the numerator in S_1^2 & S_2^2, say $S_1^2 > S_2^2$. Then

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

So that F value is always greater than or equal to 1.

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

Applications of F- test:

(i) Whether the two independent sample have been drawn from the normal population with the same variance σ^2

(ii) Whether the two independent variances are homogeneous or not.

Two random samples gave the following results.

Sample	Size	Sample mean	Sum of squares of deviations from mean
1	12	14	108
2	10	15	90

Test whether the samples came from same population.

A normal population has two parameters μ & σ^2

$\mu \rightarrow$ mean

$\sigma^2 \rightarrow$ Variance

To test whether the two samples came from same population

ie, (i) The equality of means

(ii) The Equality of Variances.

Null Hypothesis H_0: Two samples drawn from the same normal population.

ie, $\mu_1 = \mu_2$ & $\sigma_1^2 = \sigma_2^2$.

To use t-test, assumptions $\sigma_1^2 = \sigma_2^2$.

(ii) To test $\sigma_1^2 = \sigma_2^2$ (F-test)

Here $n_1 = 12$; $n_2 = 10$

$\sum (x-\bar{x})^2 = 108$

$\sum (y-\bar{y})^2 = 90.$

$$S_1^2 = \frac{\sum (x-\bar{x})^2}{n_1 - 1} = \frac{108}{11} = 9.818$$

$$S_2^2 = \frac{\sum (y-\bar{y})^2}{n_2 - 1} = \frac{90}{9} = 10$$

$S_2^2 > S_1^2$

$$F = \frac{S_2^2}{S_1^2} = \frac{10}{9.818} = 1.0185$$

Tabulalated Value at 5% level of Significance at

$(9,11) = 2.90$.

Calculated Value of $F = 1.0185$

$1.0185 < 2.90$.

$\therefore$ Accept the null hypothesis. i.e., There is no significant difference in Variances or $\sigma_1^2 = \sigma_2^2$.

$\therefore$ Both the Samples Come from the population of equal Variance.

(ii) To test $\mu_1 = \mu_2$ we use t-test

$$S^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2} = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{108+90}{12+10-2} = \frac{108+90}{20} = 9.9$$

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

t-test statistic : $t = \dfrac{\bar{x}-\bar{y}}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \dfrac{14-15}{\sqrt{9.9\left(\frac{1}{10}+\frac{1}{12}\right)}} = 0.7422$

Tabulated Value $\nu = 20$ d.f.

$$= (n_1 + n_2 - 2) = 20$$

$$t_{(0.05,\, 20)} = 2.086$$

Calculated Value of t < table Value of t at 5% level of Significance i.e., $0.7422 < 2.086$.

$\therefore$ Accept the null hypothesis i.e., $H_0: \mu_1 = \mu_2$.

$\therefore$ The two samples have been drawn from the same population.

3) Values of a Variate in two Samples are given below

Sample 1 : 5 6 8 1 12 4 3 9 6 10

Sample 2 : 2 3 6 8 1 10 2 8 – –

Test whether there is any significance difference between Sample Variances :

$n_1 = 10 \quad n_2 = 8$

H_0 : There is no significant difference between the Sample Variances.

H_1 : There is significant difference between the Sample Variances.

$$s_1^2 = \frac{1}{n_1 - 1} \sum (x_i - \bar{x})^2$$ also $$s_1^2 = \frac{\sum x_i^2}{n_1} - \left(\frac{\sum x_i}{n}\right)^2.$$

$$s_2^2 = \frac{1}{n_2 - 1} \sum (y_i - \bar{y})^2$$ $$s_2^2 = \frac{\sum y_i^2}{n_2} - \left(\frac{\sum y_i}{n}\right)^2$$

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

x	x^2	y	y^2
5	25	2	4
6	36	3	9
8	64	6	36
1	1	8	64
12	144	1	1
4	16	10	100
3	9	2	4
9	81	8	64
6	36		
10	100		
64	512	40	282

$$s_1^2 = \frac{512}{10} - (6.4)^2 = 10.24.$$

$$s_2^2 = \frac{282}{8} - 5^2 = 10.25$$

$$S_1^2 = \frac{10 \times 10.24}{9} = 11.37$$

$$S_2^2 = \frac{8 \times 10.25}{7} = 11.714.$$

Now $F = \frac{S_2^2}{S_1^2} = \frac{11.74}{11.37} = 1.03$

$d.f = (7, 9)$.

F-table Value at 5% level of significance is 3.29.

F-calculated Value < F-table Value.

$\therefore$ Accept the null hypothesis ie, There is no difference in the Sample Variances.