F-TEST:

A large number of surveys or experiments are conducted to draw conclusions about the effect of certain factors or treatments. Observations are taken pertaining to the character under study. F-test is used either for testing the hypothesis about the equality of two population variances or the equality of two or more population means. Because of this reason, it is considered to be very poplar and useful distribution and is the backbone of analysis of variance.

F-STATISTIC:

If X is a χ^2- varaite with n_1 degree of freedom and Y is an independent χ^2- varaite with n_2 degree of freedom, then F- Statistic is defined as :

$$F = \frac{X/n_1}{Y/n_2}$$

and it follows G.W Snedecor's F- distribution with (n_1,n_2) d.f.

F-TEST FOR EQUALITY OF POPULATION VARIANCES:

Let $x_1, x_2,\ldots\ldots\ldots x_{n1}$ be a random sample of size n_1 from the first normal population with variance ${\sigma_1}^2$ and $y_1, y_2,\ldots\ldots\ldots y_{n2}$ be a random sample of size n_2 from the second normal population with variance ${\sigma_2}^2$. Obviouslt the two samples rae independent. We set up the null hypothesis H_0:${\sigma_1}^2 = {\sigma_2}^2 = \sigma^2$ i.e population variances are same. In other words H_0 is that the two independent estimates of the common population variance are homogeneous i.e do not differ significantly.

Under H_0, the test statistic is given as:

$$F = \frac{{S_1}^2}{{S_2}^2} \sim F(n_1, n_2)$$

Where ${S_1}^2 = \frac{1}{n_1 - 1}\sum(x_i - \bar{x})^2$, ${S_2}^2 = \frac{1}{n_2 - 1}\sum(y_i - \bar{y})^2$.

Since F- test is based on the ratio of two variances, it is also known as variance ratio test. Also, it should be noted that the available tables of the significant values of F are for the right-tail test i.e against the alternative H$_1$:$\sigma_1{}^2 > \sigma_2{}^2$, in numerical problems we will take greater of the variance $S_1{}^2$ or $S_2{}^2$ in the numerator and adjust for the degrees of freedom accordingly.

Ex: Life expectancy in 9 regions of Brazil in 1990 and in 11 regions of Brazil in 1970 was as given in the table below:

Regions	1	2	3	4	5	6	7	8	9	10	11
Life expectancy (Yrs) 1900	42.7	43.7	34.0	39.2	46.1	48.7	49.4	45.9	55.3	-	-
Life expectancy (Yrs) 1970	54.2	50.4	44.2	49.7	55.4	57.0	58.2	56.6	61.9	57.5	53.4

Test whether the variation in life expectancy in various regions in 1900 and in 1970 is same or not.

Sol: First of all we set up the null hypothesis H$_0$:$\sigma_1{}^2 = \sigma_2{}^2$ *against* H$_1$:$\sigma_1{}^2 \neq \sigma_2{}^2$.

Here $\sum x = 405,\ \sum x^2 = 18527.78, \sum y = 598.5,\ \sum y^2 = 32799.91$

$$s_1{}^2 = \frac{1}{n_1 - 1}\left[\sum x_i{}^2 - \frac{(\sum x_i)^2}{n_1}\right]$$

$$= 37.848$$

$$s_2{}^2 = \frac{1}{n_2 - 1}\left[\sum y_i{}^2 - \frac{(\sum y_i)^2}{n_2}\right]$$

$$== 23.607$$

Since $s_1{}^2 > s_2{}^2$

The test statistic is given as:

$$F = \frac{S_1{}^2}{S_2{}^2} \sim F(n_1, n_2)$$

$$= 1.603$$

The tabulated value of F at α=0.05 and (8,10) d.f is 3.85. Since calculated F is less than the tabulated F, therefore, we accept our null hypothesis and conclude that variation in the life expectancy in various regions of Brazil in 1900 and 1970 is same.