

## 12.2: Test of Two Variances

This chapter introduces a new probability density function, the  $F$  distribution. This distribution is used for many applications including ANOVA and for testing equality across multiple means. We begin with the  $F$  distribution and the test of hypothesis of differences in variances. It is often desirable to compare two variances rather than two averages. For instance, college administrators would like two college professors grading exams to have the same variation in their grading. In order for a lid to fit a container, the variation in the lid and the container should be approximately the same. A supermarket might be interested in the variability of check-out times for two checkers. In finance, the variance is a measure of risk and thus an interesting question would be to test the hypothesis that two different investment portfolios have the same variance, the volatility.

In order to perform a  $F$  test of two variances, it is important that the following are true:

1. The populations from which the two samples are drawn are approximately normally distributed.
2. The two populations are independent of each other.

Unlike most other hypothesis tests in this book, the  $F$  test for equality of two variances is very sensitive to deviations from normality. If the two distributions are not normal, or close, the test can give a biased result for the test statistic.

Suppose we sample randomly from two independent normal populations. Let  $\sigma_1^2$  and  $\sigma_2^2$  be the unknown population variances and  $s_1^2$  and  $s_2^2$  be the sample variances. Let the sample sizes be  $n_1$  and  $n_2$ . Since we are interested in comparing the two sample variances, we use the  $F$  ratio:

$$F = \frac{\left[ \frac{(s_1)^2}{(\sigma_1)^2} \right]}{\left[ \frac{(s_2)^2}{(\sigma_2)^2} \right]} \quad (12.2.1)$$

$F$  has the distribution  $F \sim F(n_1 - 1, n_2 - 1)$

where  $n_1 - 1$  are the degrees of freedom for the numerator and  $n_2 - 1$  are the degrees of freedom for the denominator.

If the null hypothesis is  $\sigma_1^2 = \sigma_2^2$ , then the  $F$  Ratio, test statistic, becomes  $F_c = \frac{\left[ \frac{s_1^2}{\sigma_1^2} \right]}{\left[ \frac{s_2^2}{\sigma_2^2} \right]} = \frac{s_1^2}{s_2^2}$

The various forms of the hypotheses tested are:

Two-Tailed Test	One-Tailed Test	One-Tailed Test
$H_0: \sigma_1^2 = \sigma_2^2$	$H_0: \sigma_1^2 \leq \sigma_2^2$	$H_0: \sigma_1^2 \geq \sigma_2^2$
$H_1: \sigma_1^2 \neq \sigma_2^2$	$H_1: \sigma_1^2 > \sigma_2^2$	$H_1: \sigma_1^2 < \sigma_2^2$

A more general form of the null and alternative hypothesis for a two tailed test would be :

$$\begin{aligned} H_0 : \frac{\sigma_1^2}{\sigma_2^2} &= \delta_0 \\ H_a : \frac{\sigma_1^2}{\sigma_2^2} &\neq \delta_0 \end{aligned} \quad (12.2.2)$$

Where if  $\delta_0 = 1$  it is a simple test of the hypothesis that the two variances are equal. This form of the hypothesis does have the benefit of allowing for tests that are more than for simple differences and can accommodate tests for specific differences as we did for differences in means and differences in proportions. This form of the hypothesis also shows the relationship between the  $F$  distribution and the  $\chi^2$ : the  $F$  is a ratio of two chi squared distributions a distribution we saw in the last chapter. This is helpful in determining the degrees of freedom of the resultant  $F$  distribution

If the two populations have equal variances, then  $s_1^2$  and  $s_2^2$  are close in value and the test statistic,  $F_c = \frac{s_1^2}{s_2^2}$  is close to one. But if the two population variances are very different,  $s_1^2$  and  $s_2^2$  tend to be very different, too. Choosing  $s_1^2$  as the larger sample variance causes the ratio  $\frac{s_1^2}{s_2^2}$  to be greater than one. If  $s_1^2$  and  $s_2^2$  are far apart, then  $F_c = \frac{s_1^2}{s_2^2}$  is a large number.

Therefore, if  $F$  is close to one, the evidence favors the null hypothesis (the two population variances are equal). But if  $F$  is much

larger than one, then the evidence is against the null hypothesis. In essence, we are asking if the calculated  $F$  statistic, test statistic, is significantly different from one.

To determine the critical points we have to find  $F_{\alpha, df1, df2}$ . See Appendix A for the  $F$  table. This  $F$  table has values for various levels of significance from 0.1 to 0.001 designated as " $p$ " in the first column. To find the critical value choose the desired significance level and follow down and across to find the critical value at the intersection of the two different degrees of freedom. The  $F$  distribution has two different degrees of freedom, one associated with the numerator,  $df1$ , and one associated with the denominator,  $df2$  and to complicate matters the  $F$  distribution is not symmetrical and changes the degree of skewness as the degrees of freedom change. The degrees of freedom in the numerator is  $n_1 - 1$ , where  $n_1$  is the sample size for group 1, and the degrees of freedom in the denominator is  $n_2 - 1$ , where  $n_2$  is the sample size for group 2.  $F_{\alpha, df1, df2}$  will give the critical value on the upper end of the  $F$  distribution.

To find the critical value for the lower end of the distribution, reverse the degrees of freedom and divide the  $F$ -value from the table into one.

- Upper tail critical value :  $F_{\alpha, df1, df2}$
- Lower tail critical value :  $1/F_{\alpha, df2, df1}$

When the calculated value of  $F$  is between the critical values, not in the tail, we cannot reject the null hypothesis that the two variances came from a population with the same variance. If the calculated  $F$ -value is in either tail we cannot accept the null hypothesis just as we have been doing for all of the previous tests of hypothesis.

An alternative way of finding the critical values of the  $F$  distribution makes the use of the  $F$ -table easier. We note in the  $F$ -table that all the values of  $F$  are greater than one therefore the critical  $F$  value for the left hand tail will always be less than one because to find the critical value on the left tail we divide an  $F$  value into the number one as shown above. We also note that if the sample variance in the numerator of the test statistic is larger than the sample variance in the denominator, the resulting  $F$  value will be greater than one. The shorthand method for this test is thus to be sure that the larger of the two sample variances is placed in the numerator to calculate the test statistic. This will mean that only the right hand tail critical value will have to be found in the  $F$ -table.

### ✓ Exercise 12.2.1

Two college instructors are interested in whether or not there is any variation in the way they grade math exams. They each grade the same set of 30 exams. The first instructor's grades have a variance of 52.3. The second instructor's grades have a variance of 89.9. Test the claim that the first instructor's variance is smaller. (In most colleges, it is desirable for the variances of exam grades to be nearly the same among instructors.) The level of significance is 10%.

#### Answer

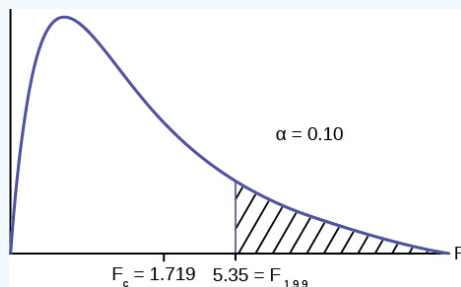
Let 1 and 2 be the subscripts that indicate the first and second instructor, respectively.

$$\begin{aligned} n_1 &= n_2 = 10 \\ H_0 : \sigma_1^2 &\geq \sigma_2^2 \text{ and } H_a : \sigma_1^2 < \sigma_2^2 \end{aligned} \quad (12.2.3)$$

**Calculate the test statistic:** By the null hypothesis ( $\sigma_1^2 \geq \sigma_2^2$ ), the  $F$  statistic is:

$$F_c = \frac{s_2^2}{s_1^2} = \frac{89.9}{52.3} = 1.719 \quad (12.2.4)$$

**Critical value for the test:**  $F_{9,9} = 5.35$  where  $n_1 - 1 = 9$  and  $n_2 - 1 = 9$ .



Make a decision: Since the calculated F value is not in the tail we cannot reject  $H_0$ .

Conclusion: With a 10% level of significance, from the data, there is insufficient evidence to conclude that the variance in grades for the first instructor is smaller.

### ? Exercise 12.2.2

The New York Choral Society divides male singers up into four categories from highest voices to lowest: Tenor1, Tenor2, Bass1, Bass2. In the table are heights of the men in the Tenor1 and Bass2 groups. One suspects that taller men will have lower voices, and that the variance of height may go up with the lower voices as well. Do we have good evidence that the variance of the heights of singers in each of these two groups (Tenor1 and Bass2) are different?

Tenor1	Bass2	Tenor 1	Bass 2	Tenor 1	Bass 2
69	72	67	72	68	67
72	75	70	74	67	70
71	67	65	70	64	70
66	75	72	66		69
76	74	70	68		72
74	72	68	75		71
71	72	64	68		74
66	74	73	70		75
68	72	66	72		

#### Answer

The histograms are not as normal as one might like. Plot them to verify. However, we proceed with the test in any case.

Subscripts: T1 = tenor 1 and B2 = bass 2

The standard deviations of the samples  $s_{T1} = 3.3302$  and  $s_{B2} = 2.7208$ .

The hypotheses are

$H_0 : \sigma_{T1}^2 = \sigma_{B2}^2$  and  $H_a : \sigma_{T1}^2 \neq \sigma_{B2}^2$  (two tailed test)

The  $F$  statistic is 1.4894 with 20 and 25 degrees of freedom.

The  $p$ -value is 0.3430. If we assume alpha is 0.05, then we cannot reject the null hypothesis.

We have no good evidence from the data that the heights of Tenor1 and Bass2 singers have different variances (despite there being a significant difference in mean heights of about 2.5 inches.)

This page titled [12.2: Test of Two Variances](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [OpenStax](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.

- [12.1: Test of Two Variances](#) has no license indicated.