12 F DISTRIBUTION AND ONE-WAY ANOVA

Figure 12.1 One-way ANOVA is used to measure information from several groups. (credit: modification of work "Magazine Stack" by thebittenword.com/ Flickr, CC BY 2.0)

Introduction

Many statistical applications in psychology, social science, business administration, and the natural sciences involve several groups. For example, an environmentalist is interested in knowing if the average amount of pollution varies in several bodies of water. A sociologist is interested in knowing if the amount of income a person earns varies according to their upbringing. A consumer looking for a new car might compare the average gas mileage of several models.

For hypothesis tests comparing averages among more than two groups, statisticians have developed a method called "**Analysis of Variance**" (abbreviated ANOVA). In this chapter, you will study the simplest form of ANOVA called single factor or **one-way ANOVA**. You will also study the *F* distribution, used for one-way ANOVA, and the test for differences between two variances. This is just a very brief overview of one-way ANOVA. One-Way ANOVA, as it is presented here, relies heavily on a calculator or computer.

12.1 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 σ_1^2 and σ_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:

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

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: $\sigma_1{}^2 < \sigma_2{}^2$

Table 12.1

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

$$H_0 : \frac{\sigma_1{}^2}{\sigma_2{}^2} = \delta_0$$

$$H_a : \frac{\sigma_1{}^2}{\sigma_2{}^2} \neq \delta_0$$

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 χ^2 : the F is a ratio of two chi squared distributions a distribution we saw in the The Chi-Square Distribution. 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.

EXAMPLE 12.1

Problem

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 10 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%.

Solution

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

$n_1 = n_2 = 10$.

H_0: $\sigma_1^2 \geq \sigma_2^2$ and H_a: $\sigma_1^2 < \sigma_2^2$

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

$$F_c = \frac{{s_2}^2}{{s_1}^2} = \frac{89.9}{52.3} = 1.719$$

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

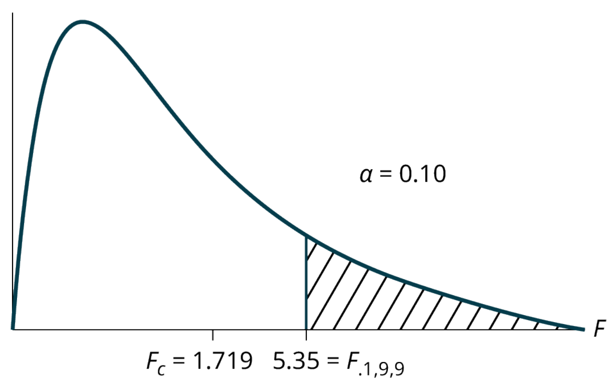


Figure 12.2

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.

TRY IT 12.1

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	Bass 2	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		

Table 12.2

12.2 One-Way ANOVA

The purpose of a one-way ANOVA test is to determine the existence of a statistically significant difference among several group means. The test actually uses **variances** to help determine if the means are equal or not. In order to perform a one-way ANOVA test, there are five basic **assumptions** to be fulfilled:

1. Each population from which a sample is taken is assumed to be normal.
2. All samples are randomly selected and independent.
3. The populations are assumed to have **equal standard deviations (or variances)**.
4. The factor is a categorical variable.
5. The response is a numerical variable.

The Null and Alternative Hypotheses

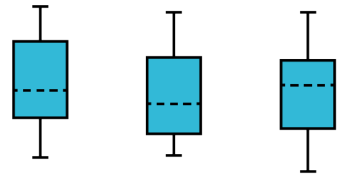
The null hypothesis is simply that all the group population means are the same. The alternative hypothesis is that at least one pair of means is different. For example, if there are *k* groups:

$H_0 : \mu_1 = \mu_2 = \mu_3 =. . . \mu_k$

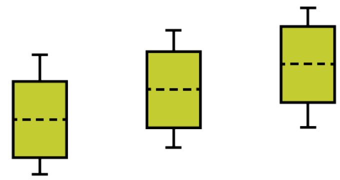
H_a: At least two of the group means $\mu_1, \mu_2, \mu_3, . . . , \mu_k$ are not equal. That is, $\mu_i \neq \mu_j$ for some $i \neq j$.

The graphs, a set of box plots representing the distribution of values with the group means indicated by a horizontal line through the box, help in the understanding of the hypothesis test. In the first graph (red box plots), H_0: $\mu_1 = \mu_2 = \mu_3$ and the three populations have the same distribution if the null hypothesis is true. The variance of the combined data is approximately the same as the variance of each of the populations.

If the null hypothesis is false, then the variance of the combined data is larger which is caused by the different means as shown in the second graph (green box plots).



(a)



(b)

Figure 12.3 (a) H_0 is true. All means are the same; the differences are due to random variation. (b) H_0 is not true. All means are not the same; the differences are too large to be due to random variation.

12.3 The F Distribution and the F-Ratio

The distribution used for the hypothesis test is a new one. It is called the ***F* distribution**, invented by George Snedecor but named in honor of Sir Ronald Fisher, an English statistician. The *F* statistic is a ratio (a fraction). There are two sets of degrees of freedom; one for the numerator and one for the denominator.

For example, if *F* follows an *F* distribution and the number of degrees of freedom for the numerator is four, and the number of degrees of freedom for the denominator is ten, then $F \sim F_{4,10}$.

To calculate the ***F* ratio**, two estimates of the variance are made.

1. **Variance between samples:** An estimate of σ^2 that is the variance of the sample means multiplied by *n* (when the sample sizes are the same.). If the samples are different sizes, the variance between samples is weighted to account for the different sample sizes. The variance is also called **variation due to treatment or explained variation.**
2. **Variance within samples:** An estimate of σ^2 that is the average of the sample variances (also known as a pooled variance). When the sample sizes are different, the variance within samples is weighted. The variance is also called the **variation due to error or unexplained variation.**

- SS_{between} = the **sum of squares** that represents the variation among the different samples
- SS_{within} = the sum of squares that represents the variation within samples that is due to chance.

To find a "sum of squares" means to add together squared quantities that, in some cases, may be weighted. We used sum of squares to calculate the sample variance and the sample standard deviation in Descriptive Statistics.

MS means "**mean square**." MS_{between} is the variance between groups, and MS_{within} is the variance within groups.

Calculation of Sum of Squares and Mean Square

- k = the number of different groups
- n_j = the size of the j^{th} group
- s_j = the sum of the values in the j^{th} group
- n = total number of all the values combined (total sample size: $\sum n_j$)
- x = one value: $\sum x = \sum s_j$
- Sum of squares of all values from every group combined: $\sum x^2$
- Between group variability: $SS_{\text{total}} = \sum x^2 - \frac{\left(\sum x^2\right)}{n}$
- Total sum of squares: $\sum x^2 - \frac{\left(\sum x\right)^2}{n}$
- Explained variation: sum of squares representing variation among the different samples: $SS_{\text{between}} = \sum \left[\frac{(s_j)^2}{n_j}\right] - \frac{\left(\sum s_j\right)^2}{n}$
- Unexplained variation: sum of squares representing variation within samples due to chance: $SS_{\text{within}} = SS_{\text{total}} - SS_{\text{between}}$
- *df*'s for different groups (*df*'s for the numerator): $df = k - 1$

- Equation for errors within samples (df's for the denominator): $df_{\text{within}} = n - k$
- Mean square (variance estimate) explained by the different groups: $MS_{\text{between}} = \frac{SS_{\text{between}}}{df_{\text{between}}}$
- Mean square (variance estimate) that is due to chance (unexplained): $MS_{\text{within}} = \frac{SS_{\text{within}}}{df_{\text{within}}}$

MS_{between} and MS_{within} can be written as follows:

- $MS_{\text{between}} = \frac{SS_{\text{between}}}{df_{\text{between}}} = \frac{SS_{\text{between}}}{k-1}$
- $MS_{within} = \frac{SS_{within}}{df_{within}} = \frac{SS_{within}}{n-k}$

The one-way ANOVA test depends on the fact that MS_{between} can be influenced by population differences among means of the several groups. Since MS_{within} compares values of each group to its own group mean, the fact that group means might be different does not affect MS_{within}.

The null hypothesis says that all groups are samples from populations having the same normal distribution. The alternate hypothesis says that at least two of the sample groups come from populations with different normal distributions. If the null hypothesis is true, MS_{between} and MS_{within} should both estimate the same value.

NOTE

The null hypothesis says that all the group population means are equal. The hypothesis of equal means implies that the populations have the same normal distribution, because it is assumed that the populations are normal and that they have equal variances.

_F_-Ratio or _F_ Statistic

$$F = \frac{MS_{\text{between}}}{MS_{\text{within}}}$$

If MS_{between} and MS_{within} estimate the same value (following the belief that H_0 is true), then the F-ratio should be approximately equal to one. Mostly, just sampling errors would contribute to variations away from one. As it turns out, MS_{between} consists of the population variance plus a variance produced from the differences between the samples. MS_{within} is an estimate of the population variance. Since variances are always positive, if the null hypothesis is false, MS_{between} will generally be larger than MS_{within}.Then the F-ratio will be larger than one. However, if the population effect is small, it is not unlikely that MS_{within} will be larger in a given sample.

The foregoing calculations were done with groups of different sizes. If the groups are the same size, the calculations simplify somewhat and the F-ratio can be written as:

_F_-Ratio Formula when the groups are the same size

$$F = \frac{n \cdot s_{\overline{x}}{}^2}{s^2{}_{\text{pooled}}}$$

where ...

- n = the sample size
- $df_{\text{numerator}} = k - 1$
- $df_{\text{denominator}} = n - k$
- s^2 pooled = the mean of the sample variances (pooled variance)
- $s_{\overline{x}}{}^2$ = the variance of the sample means

Data are typically put into a table for easy viewing. One-Way ANOVA results are often displayed in this manner by computer software.

Source of variation	Sum of squares (SS)	Degrees of freedom (df)	Mean square (MS)	F
Factor (Between)	SS(Factor)	$k - 1$	MS(Factor) = SS(Factor)/(k – 1)	F = MS(Factor)/MS(Error)
Error (Within)	SS(Error)	$n - k$	MS(Error) = SS(Error)/(n – k)	
Total	SS(Total)	$n - 1$		

Table 12.3

EXAMPLE 12.2

Three different diet plans are to be tested for mean weight loss. The entries in the table are the weight losses for the different plans. The one-way ANOVA results are shown in Table 12.4.

Plan 1: $n_1 = 4$	Plan 2: $n_2 = 3$	Plan 3: $n_3 = 3$
5	3.5	8
4.5	7	4
4		3.5
3	4.5	

Table 12.4

$s_1 = 16.5$, $s_2 = 15$, $s_3 = 15.5$

Following are the calculations needed to fill in the one-way ANOVA table. The table is used to conduct a hypothesis test.

$$SS(between) = \sum \left[\frac{(s_j)^2}{n_j}\right] - \frac{\left(\sum s_j\right)^2}{n}$$

$$= \frac{s_1^2}{4} + \frac{s_2^2}{3} + \frac{s_3^2}{3} - \frac{(s_1 + s_2 + s_3)^2}{10}$$

where $n_1 = 4$, $n_2 = 3$, $n_3 = 3$ and $n = n_1 + n_2 + n_3 = 10$

$$= \frac{(16.5)^2}{4} + \frac{(15)^2}{3} + \frac{(15.5)^2}{3} - \frac{(16.5 + 15 + 15.5)^2}{10}$$

$$SS(between) = 2.2458$$

$$S(total) = \sum x^2 - \frac{\left(\sum x\right)^2}{n}$$

$$= \left(5^2 + 4.5^2 + 4^2 + 3^2 + 3.5^2 + 7^2 + 4.5^2 + 8^2 + 4^2 + 3.5^2\right)$$

$$- \frac{(5 + 4.5 + 4 + 3 + 3.5 + 7 + 4.5 + 8 + 4 + 3.5)^2}{10}$$

$$= 244 - \frac{47^2}{10} = 244 - 220.9$$

$$SS(total) = 23.1$$

$$SS(within) = SS(total) - SS(between)$$

$$= 23.1 - 2.2458$$
$$SS(within) = 20.8542$$

Source of variation	Sum of squares (*SS*)	Degrees of freedom (*df*)	Mean square (*MS*)	*F*
Factor (Between)	*SS*(Factor) = *SS*(Between) = 2.2458	*k* – 1 = 3 groups – 1 = 2	*MS*(Factor) = *SS*(Factor)/(*k* – 1) = 2.2458/2 = 1.1229	*F* = *MS*(Factor)/*MS*(Error) = 1.1229/2.9792 = 0.3769
Error (Within)	*SS*(Error) = *SS*(Within) = 20.8542	*n* – *k* = 10 total data – 3 groups = 7	*MS*(Error) = *SS*(Error)/(*n* – *k*) = 20.8542/7 = 2.9792	
Total	*SS*(Total) = 2.2458 + 20.8542 = 23.1	*n* – 1 = 10 total data – 1 = 9		

Table 12.5

TRY IT 12.2

As part of an experiment to see how different types of soil cover would affect slicing tomato production, Marist College students grew tomato plants under different soil cover conditions. Groups of three plants each had one of the following treatments

- bare soil
- a commercial ground cover
- black plastic
- straw
- compost

All plants grew under the same conditions and were the same variety. Students recorded the weight (in grams) of tomatoes produced by each of the n = 15 plants:

Bare: $n_1 = 3$	Ground Cover: $n_2 = 3$	Plastic: $n_3 = 3$	Straw: $n_4 = 3$	Compost: $n_5 = 3$
2,625	5,348	6,583	7,285	6,277
2,997	5,682	8,560	6,897	7,818
4,915	5,482	3,830	9,230	8,677

Table 12.6

Create the one-way ANOVA table.

The one-way ANOVA hypothesis test is always right-tailed because larger *F*-values are way out in the right tail of the *F*-distribution curve and tend to make us reject H_0.

EXAMPLE 12.3

Problem

Let's return to the slicing tomato exercise in Try It 12.2. The means of the tomato yields under the five mulching conditions are represented by μ_1, μ_2, μ_3, μ_4, μ_5. We will conduct a hypothesis test to determine if all means are the same or at least one is different. Using a significance level of 5%, test the null hypothesis that there is no difference in mean yields among the five groups against the alternative hypothesis that at least one mean is different from the rest.

Solution

The null and alternative hypotheses are:

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

H_a: $\mu_i \neq \mu_j$ some $i \neq j$

The one-way ANOVA results are shown in Table 12.7

Source of variation	Sum of squares (*SS*)	Degrees of freedom (*df*)	Mean square (*MS*)	*F*
Factor (Between)	36,648,561	5 – 1 = 4	$\frac{36{,}648{,}561}{4} = 9{,}162{,}140$	$\frac{9{,}162{,}140}{2{,}044{,}672.6} = 4.4810$
Error (Within)	20,446,726	15 – 5 = 10	$\frac{20{,}446{,}726}{10} = 2{,}044{,}672.6$	
Total	57,095,287	15 – 1 = 14		

Table 12.7

Distribution for the test: $F_{4,10}$

$df(num) = 5 - 1 = 4$

$df(denom) = 15 - 5 = 10$

Test statistic: $F = 4.4810$

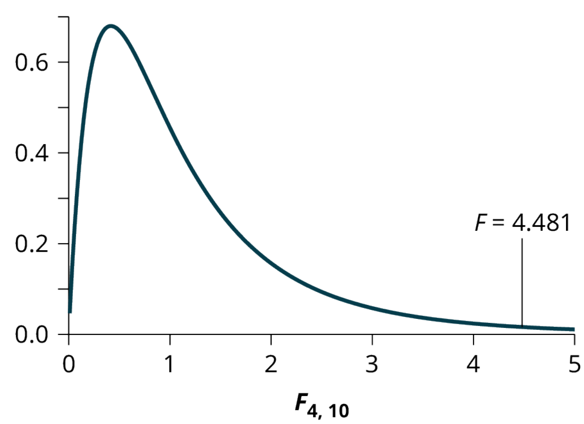


Figure 12.4

Probability Statement: p-value = $P(F > 4.481) = 0.0248$.

Compare α and the p-value: $\alpha = 0.05$, p-value = 0.0248

Make a decision: Since $\alpha > p$-value, we cannot accept H_0.

Conclusion: At the 5% significance level, we have reasonably strong evidence that differences in mean yields for slicing tomato plants grown under different mulching conditions are unlikely to be due to chance alone. We may conclude that at least some of mulches led to different mean yields.

TRY IT 12.3

There are multiple variants of the virus that causes COVID-19. The length of hospital stays for patients afflicted with various strains of COVID-19 is shown in Table 12.8.

Delta Strain	Omicron Strain	Alpha Strain	Gamma Strain	Beta Strain
13.9	11.7	18.2	16.9	9.3
14.9	15.1	14.6	12.8	15.8
16.8	9.9	10.1	11.2	16.4

Table 12.8

Test whether the mean length of hospital stay is the same or different for the various strains of COVID-19. Construct the ANOVA table, find the *p*-value, and state your conclusion. Use a 5% significance level.

EXAMPLE 12.4

Four sororities took a random sample of sisters regarding their grade means for the past term. The results are shown in Table 12.9.

Sorority 1	Sorority 2	Sorority 3	Sorority 4
2.17	2.63	2.63	3.79
1.85	1.77	3.78	3.45
2.83	3.25	4.00	3.08
1.69	1.86	2.55	2.26
3.33	2.21	2.45	3.18

Table 12.9 Mean grades for four sororities

Problem

Using a significance level of 1%, is there a difference in mean grades among the sororities?

Solution

Let μ_1, μ_2, μ_3, μ_4 be the population means of the sororities. Remember that the null hypothesis claims that the sorority groups are from the same normal distribution. The alternate hypothesis says that at least two of the sorority groups come from populations with different normal distributions. Notice that the four sample sizes are each five.

NOTE

This is an example of a **balanced design**, because each factor (i.e., sorority) has the same number of observations.

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

H_a: Not all of the means $\mu_1, \mu_2, \mu_3, \mu_4$ are equal.

Distribution for the test: $F_{3,16}$

where k = 4 groups and n = 20 samples in total

$df(num)= k - 1 = 4 - 1 = 3$

$df(denom) = n - k = 20 - 4 = 16$

Calculate the test statistic: $F = 2.23$

Graph:

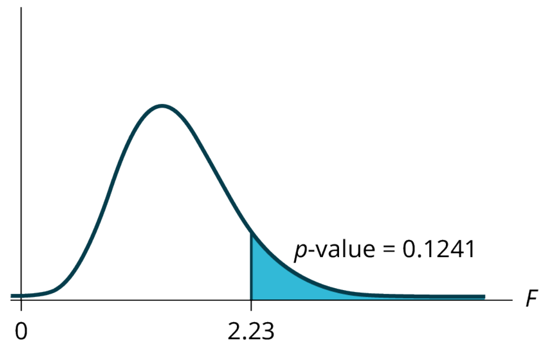


Figure 12.5

Probability statement: p-value = $P(F > 2.23) = 0.1241$

Compare α and the p-value: $\alpha = 0.01$
p-value = 0.1241
$\alpha < p$-value

Make a decision: Since $\alpha < p$-value, you cannot reject H_0.

Conclusion: There is not sufficient evidence to conclude that there is a difference among the mean grades for the sororities.

TRY IT 12.4

Four sports teams took a random sample of players regarding their GPAs for the last year. The results are shown in Table 12.10.

Basketball	Baseball	Hockey	Lacrosse
3.6	2.1	4.0	2.0
2.9	2.6	2.0	3.6
2.5	3.9	2.6	3.9
3.3	3.1	3.2	2.7
3.8	3.4	3.2	2.5

Table 12.10 GPAs for four sports teams

Use a significance level of 5%, and determine if there is a difference in GPA among the teams.

EXAMPLE 12.5

A fourth grade class is studying the environment. One of the assignments is to grow bean plants in different soils. Tommy chose to grow his bean plants in soil found outside his classroom mixed with dryer lint. Tara chose to grow her bean plants in potting soil bought at the local nursery. Nick chose to grow his bean plants in soil from his mother's garden. No chemicals were used on the plants, only water. They were grown inside the classroom next to a large window. Each child grew five plants. At the end of the growing period, each plant was measured, producing the data (in inches) in Table 12.11.

Tommy's plants	Tara's plants	Nick's plants
24	25	23
21	31	27
23	23	22
30	20	30
23	28	20

Table 12.11

Problem

Does it appear that the three media in which the bean plants were grown produce the same mean height? Test at a 3% level of significance.

Solution

This time, we will perform the calculations that lead to the *F'* statistic. Notice that each group has the same number of plants, so we will use the formula $F' = \frac{n \cdot s_{\overline{x}}{}^2}{s^2{}_{\text{pooled}}}$.

First, calculate the sample mean and sample variance of each group.

	Tommy's plants	Tara's plants	Nick's plants
Sample mean	24.2	25.4	24.4
Sample variance	11.7	18.3	16.3

Table 12.12

Next, calculate the variance of the three group means (Calculate the variance of 24.2, 25.4, and 24.4). **Variance of the group means = 0.413** = $s_{\overline{x}}{}^2$

Then $MS_{between} = ns_{\overline{x}}{}^2$ = (5)(0.413) where $n = 5$ is the sample size (number of plants each child grew).

Calculate the mean of the three sample variances (Calculate the mean of 11.7, 18.3, and 16.3). **Mean of the sample variances = 15.433** = s^2 ***pooled***

Then $MS_{within} = s^2{}_{pooled} = 15.433$.

The *F* statistic (or *F* ratio) is $F = \frac{MS_{\text{between}}}{MS_{\text{within}}} = \frac{ns_{\overline{x}}{}^2}{s^2{}_{pooled}} = \frac{(5)(0.413)}{15.433} = 0.134$

The *dfs* for the numerator = the number of groups – 1 = 3 – 1 = 2.

The *dfs* for the denominator = the total number of samples – the number of groups = 15 – 3 = 12

The distribution for the test is $F_{2,12}$ and the *F* statistic is $F = 0.134$

The *p*-value is $P(F > 0.134) = 0.8759$.

Decision: Since $\alpha = 0.03$ and the *p*-value = 0.8759, then you cannot reject H_0. (Why?)

Conclusion: With a 3% level of significance, from the sample data, the evidence is not sufficient to conclude that the mean heights of the bean plants are different.

TRY IT 12.5

Another fourth grader also grew bean plants, but this time in a jelly-like mass. The heights were (in inches) 24, 28, 25, 30, and 32. Do a one-way ANOVA test on the four groups. Are the heights of the bean plants different? Use the same method as shown in Example 12.5.

Notation

The notation for the *F* distribution is $F \sim F_{df(num),df(denom)}$

where $df(num) = df_{\text{between}}$ and $df(denom) = df_{\text{within}}$

The mean for the *F* distribution is $\mu = \frac{df(num)}{df(denom) - 2}$

12.4 Facts About the F Distribution

Here are some facts about the *F* distribution.

1. The curve is not symmetrical but skewed to the right.
2. There is a different curve for each set of degrees of freedom.
3. The *F* statistic is greater than or equal to zero.
4. As the degrees of freedom for the numerator and for the denominator get larger, the curve approximates the normal as can be seen in the two figures below. Figure (b) with more degrees of freedom is more closely approaching the normal distribution, but remember that the *F* cannot ever be less than zero so the distribution does not have a tail that goes to infinity on the left as the normal distribution does.
5. Other uses for the *F* distribution include comparing two variances and two-way Analysis of Variance. Two-Way Analysis is beyond the scope of this chapter.

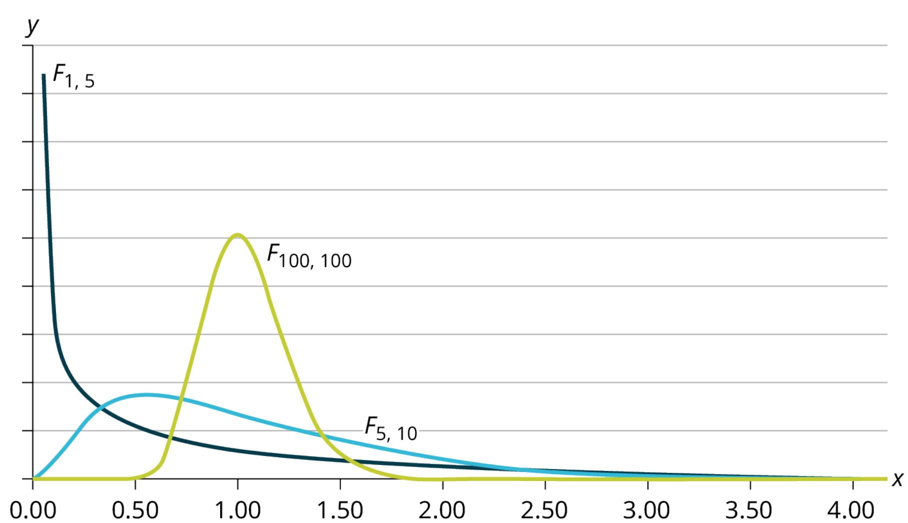


Figure 12.6