

11.4: F-Tests in One-Way ANOVA

Learning Objectives

- To understand how to use an F -test to judge whether several population means are all equal

In Chapter 9, we saw how to compare two population means μ_1 and μ_2 .

In this section we will learn to compare three or more population means at the same time, which is often of interest in practical applications. For example, an administrator at a university may be interested in knowing whether student grade point averages are the same for different majors. In another example, an oncologist may be interested in knowing whether patients with the same type of cancer have the same average survival times under several different competing cancer treatments.

In general, suppose there are K normal populations with possibly different means, $\mu_1, \mu_2, \dots, \mu_K$, but all with the same variance σ^2 . The study question is whether all the K population means are the same. We formulate this question as the test of hypotheses

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_K$$

vs.

$$H_a : \text{not all } K \text{ population means are equal}$$

To perform the test K independent random samples are taken from the K normal populations. The K sample means, the K sample variances, and the K sample sizes are summarized in the table:

Population	Sample Size	Sample Mean	Sample Variance
1	n_1	\bar{x}_1	s_1^2
2	n_2	\bar{x}_2	s_2^2
\vdots	\vdots	\vdots	\vdots
K	n_K	\bar{x}_K	s_K^2

Define the following quantities:

Definitions

The *combined sample size*:

$$n = n_1 + n_2 + \dots + n_K$$

The *mean of the combined sample* of all n observations:

$$\bar{x} = \frac{\sum x}{n} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + \dots + n_K \bar{x}_K}{n}$$

The *mean square for treatment*:

$$MST = \frac{n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + \dots + n_K(\bar{x}_K - \bar{x})^2}{K - 1}$$

The *mean square for error*:

$$MSE = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_K - 1)s_K^2}{n - K}$$

MST can be thought of as the variance between the K individual independent random samples and MSE as the variance within the samples. This is the reason for the name “analysis of variance,” universally abbreviated ANOVA. The adjective “one-way” has to do with the fact that the sampling scheme is the simplest possible, that of taking one random sample from each population under consideration. If the means of the K populations are all the same then the two quantities MST and MSE should be close to the

same, so the null hypothesis will be rejected if the ratio of these two quantities is significantly greater than 1. This yields the following test statistic and methods and conditions for its use.

Test Statistic for Testing the Null Hypothesis that K Population Means Are Equal

$$F = \frac{MST}{MSE}$$

If the K populations are normally distributed with a common variance and if $H_0 : \mu_1 = \mu_2 = \cdots = \mu_K$ is true then under independent random sampling F approximately follows an F -distribution with degrees of freedom $df_1 = K - 1$ and $df_2 = n - K$.

The test is right-tailed: H_0 is rejected at level of significance α if $F \geq F_\alpha$.

As always the test is performed using the usual five-step procedure.

Example 11.4.1

The average of grade point averages (GPAs) of college courses in a specific major is a measure of difficulty of the major. An educator wishes to conduct a study to find out whether the difficulty levels of different majors are the same. For such a study, a random sample of major grade point averages (GPA) of 11 graduating seniors at a large university is selected for each of the four majors mathematics, English, education, and biology. The data are given in Table 11.4.1. Test, at the 5% level of significance, whether the data contain sufficient evidence to conclude that there are differences among the average major GPAs of these four majors.

Table 11.4.1: Difficulty Levels of College Majors

Mathematics	English	Education	Biology
2.59	3.64	4.00	2.78
3.13	3.19	3.59	3.51
2.97	3.15	2.80	2.65
2.50	3.78	2.39	3.16
2.53	3.03	3.47	2.94
3.29	2.61	3.59	2.32
2.53	3.20	3.74	2.58
3.17	3.30	3.77	3.21
2.70	3.54	3.13	3.23
3.88	3.25	3.00	3.57
2.64	4.00	3.47	3.22

Solution

- **Step 1.** The test of hypotheses is

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

vs.

$$H_a : \text{not all four population means are equal @ } \alpha = 0.05$$

- **Step 2.** The test statistic is $F = MST/MSE$ with (since $n = 44$ and $K = 4$) degrees of freedom $df_1 = K - 1 = 4 - 1 = 3$ and $df_2 = n - K = 44 - 4 = 40$.
- **Step 3.** If we index the population of mathematics majors by 1, English majors by 2, education majors by 3, and biology majors by 4, then the sample sizes, sample means, and sample variances of the four samples in Table 11.4.1 are summarized (after rounding for simplicity) by:

Major	Sample Size	Sample Mean	Sample Variance
Mathematics	$n_1 = 11$	$\bar{x}_1 = 2.90$	$s_1^2 = 0.188$
English	$n_2 = 11$	$\bar{x}_2 = 3.34$	$s_2^2 = 0.148$
Education	$n_3 = 11$	$\bar{x}_3 = 3.36$	$s_3^2 = 0.229$
Biology	$n_4 = 11$	$\bar{x}_4 = 3.02$	$s_4^2 = 0.157$

The average of all 44 observations is (after rounding for simplicity) $\bar{x} = 3.15$. We compute (rounding for simplicity)

$$\begin{aligned}
 MST &= \frac{11(2.90 - 3.15)^2 + 11(3.34 - 3.15)^2 + 11(3.36 - 3.15)^2 + 11(3.02 - 3.15)^2}{4 - 1} \\
 &= \frac{1.7556}{3} \\
 &= 0.585
 \end{aligned}$$

and

$$\begin{aligned}
 MSE &= \frac{(11 - 1)(0.188) + (11 - 1)(0.148) + (11 - 1)(0.229) + (11 - 1)(0.157)}{44 - 4} \\
 &= \frac{7.22}{40} \\
 &= 0.181
 \end{aligned}$$

so that

$$F = \frac{MST}{MSE} = \frac{0.585}{0.181} = 3.232$$

- **Step 4.** The test is right-tailed. The single critical value is (since $df_1 = 3$ and $df_2 = 40$) $F_\alpha = F_{0.05} = 2.84$. Thus the rejection region is $[2.84, \infty)$, as illustrated in Figure 11.4.1.

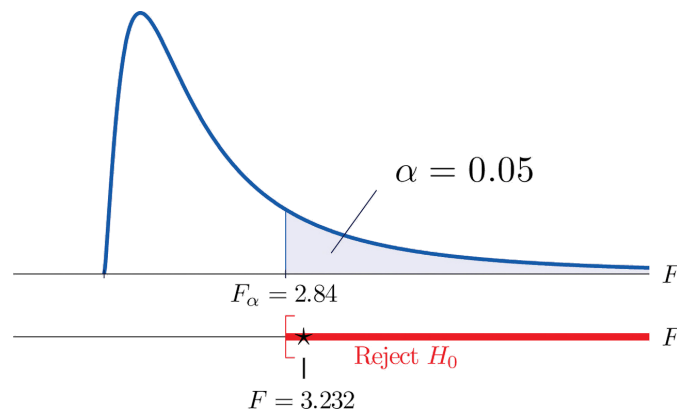


Figure 11.4.1: Rejection Region

- **Step 5.** Since $F = 3.232 > 2.84$, we reject H_0 . The data provide sufficient evidence, at the 5% level of significance, to conclude that the averages of major GPAs for the four majors considered are not all equal.

✓ Example 11.4.2: Mice Survival Times

A research laboratory developed two treatments which are believed to have the potential of prolonging the survival times of patients with an acute form of thymic leukemia. To evaluate the potential treatment effects 33 laboratory mice with thymic leukemia were randomly divided into three groups. One group received Treatment 1, one received Treatment 2, and the third was observed as a control group. The survival times of these mice are given in Table 11.4.2 Test, at the 1% level of

significance, whether these data provide sufficient evidence to confirm the belief that at least one of the two treatments affects the average survival time of mice with thymic leukemia.

Table 11.4.2 Mice Survival Times in Days

Treatment 1		Treatment 2	Control
71	75	77	81
72	73	67	79
75	72	79	73
80	65	78	71
60	63	81	75
65	69	72	84
63	64	71	77
78	71	84	67
		91	

Solution

- **Step 1.** The test of hypotheses is

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

vs.

$$H_a : \text{not all three population means are equal @ } \alpha = 0.01$$

- **Step 2.** The test statistic is $F = \frac{MST}{MSE}$ with (since $n = 33$ and $K = 3$) degrees of freedom $df_1 = K - 1 = 3 - 1 = 2$ and $df_2 = n - K = 33 - 3 = 30$.
- **Step 3.** If we index the population of mice receiving Treatment 1 by 1, Treatment 2 by 2, and no treatment by 3, then the sample sizes, sample means, and sample variances of the three samples in Table 11.4.2 are summarized (after rounding for simplicity) by:

Table 11.4.2: Mice Survival Times in Days

Group	Sample Size	Sample Mean	Sample Variance
Treatment 1	$n_1 = 16$	$\bar{x}_1 = 69.75$	$s_1^2 = 34.47$
Treatment 2	$n_2 = 9$	$\bar{x}_2 = 77.78$	$s_2^2 = 52.69$
Control	$n_3 = 8$	$\bar{x}_3 = 75.88$	$s_3^2 = 30.69$

The average of all 33 observations is (after rounding for simplicity) $\bar{x} = 73.42$. We compute (rounding for simplicity)

$$\begin{aligned}
 MST &= \frac{16(69.75 - 73.42)^2 + 9(77.78 - 73.42)^2 + 8(75.88 - 73.42)^2}{31} \\
 &= \frac{434.63}{2} \\
 &= 217.50
 \end{aligned}$$

and

$$\begin{aligned}
 MSE &= \frac{(16 - 1)(34.47) + (9 - 1)(52.69) + (8 - 1)(30.69)}{33 - 3} \\
 &= \frac{1153.4}{30} \\
 &= 38.45
 \end{aligned}$$

so that

$$F = \frac{MST}{MSE} = \frac{217.50}{38.45} = 5.65$$

- **Step 4.** The test is right-tailed. The single critical value is $F_{\alpha} = F_{0.01} = 5.39$. Thus the rejection region is $[5.39, \infty)$, as illustrated in Figure 11.4.2

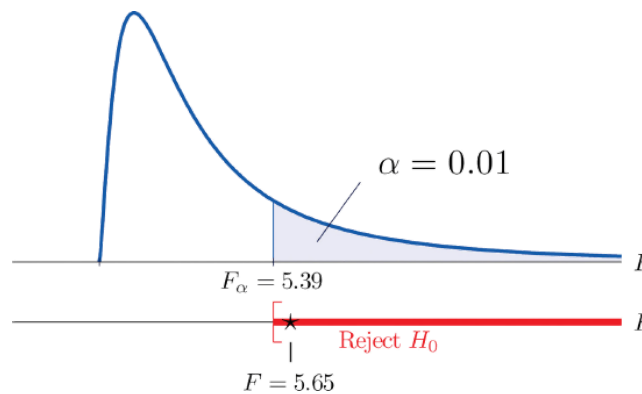


Figure 11.4.2: Rejection Region

- **Step 5.** Since $F = 5.65 > 5.39$, we reject H_0 . The data provide sufficient evidence, at the 1% level of significance, to conclude that a treatment effect exists at least for one of the two treatments in increasing the mean survival time of mice with thymic leukemia.

It is important to note that, if the null hypothesis of equal population means is rejected, the statistical implication is that not all population means are equal. It does not however tell which population mean is different from which. The inference about where the suggested difference lies is most frequently made by a follow-up study.

Key Takeaway

- An F -test can be used to evaluate the hypothesis that the means of several normal populations, all with the same standard deviation, are identical.

This page titled [11.4: F-Tests in One-Way ANOVA](#) is shared under a [CC BY-NC-SA 3.0](#) license and was authored, remixed, and/or curated by [Anonymous](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.