

## Chapter 548

# One-Way Analysis of Variance F-Tests using Effect Size

---

### Introduction

A common task in research is to compare the averages of two or more populations (groups). We might want to compare the income level of two regions, the nitrogen content of three lakes, or the effectiveness of four drugs. The one-way analysis of variance compares the means of two or more groups to determine if at least one mean is different from the others. The  $F$  test is used to determine statistical significance.  $F$  tests are non-directional in that the null hypothesis specifies that all means are equal and the alternative hypothesis simply states that at least one mean is different.

The methods described here are usually applied to the one-way experimental design. This design is an extension of the design used for the two-sample  $t$  test. Instead of two groups, there are three or more groups. With careful modifications, this procedure may be used to test interaction terms as well.

In our more advanced one-way ANOVA procedure, you are required to enter hypothesized means and variances. This simplified procedure only requires the input of an *effect size*, usually  $f$ , as proposed by Cohen (1988).

---

### Assumptions

Using the  $F$  test requires certain assumptions. One reason for the popularity of the  $F$  test is its robustness in the face of assumption violation. However, if an assumption is not even approximately met, the significance levels and the power of the  $F$  test are invalidated. Unfortunately, in practice it often happens that several assumptions are not met. This makes matters even worse. Hence, steps should be taken to check the assumptions before important decisions are made.

The assumptions of the one-way analysis of variance are:

1. The data are continuous (not discrete).
2. The data follow the normal probability distribution. Each group is normally distributed about the group mean.
3. The variances of the populations are equal.
4. The groups are independent. There is no relationship among the individuals in one group as compared to another.
5. Each group is a simple random sample from its population. Each individual in the population has an equal probability of being selected in the sample.

## Technical Details for the One-Way ANOVA

Suppose  $k$  groups each have a normal distribution and equal means ( $\mu_1 = \mu_2 = \dots = \mu_k$ ). Let  $n_1 = n_2 = \dots = n_k$  denote the number of subjects in each group and let  $N$  denote the total sample size of all groups. Let  $\bar{\mu}_w$  denote the weighted mean of all groups. That is

$$\bar{\mu}_w = \sum_{i=1}^k \left( \frac{n_i}{N} \right) \mu_i$$

Let  $\sigma$  denote a common standard deviation of all groups.

Given the above terminology, the ratio of the mean square between groups to the mean square within groups follows a central  $F$  distribution with two parameters matching the degrees of freedom of the numerator mean square and the denominator mean square. When the null hypothesis of mean equality is rejected, the above ratio has a noncentral  $F$  distribution which adds a noncentrality parameter,  $\lambda$ . This parameter is calculated as

$$\lambda = N \frac{\sigma_m^2}{\sigma^2}$$

where

$$\sigma_m = \sqrt{\frac{\sum_{i=1}^k n_i (\mu_i - \bar{\mu}_w)^2}{N}}$$

Cohen (1988) proposed a further substitution of an *effect size* parameter  $f$  which is defined as

$$f = \sqrt{\frac{\sigma_m^2}{\sigma^2}}$$

Using this parameter, the noncentrality parameter simplifies to

$$\lambda = N f^2$$

Cohen (1988, 285-287) proposed the following interpretation of  $f$ :  $f = 0.1$  is a **small** effect,  $f = 0.25$  is a **medium** effect, and  $f = 0.4$  is a **large** effect.

Cohen (1988) also referenced another effect size parameter which he named  $\eta^2$  (*eta-squared*). This parameter is defined as

$$\eta^2 = \frac{\sigma_m^2}{\sigma_m^2 + \sigma^2} = \frac{f^2}{1 + f^2}$$

$\eta^2$  is the proportion of the total variation in the dependent variable that is attributable to the groups.

---

## Power Calculations for One-Way ANOVA

The calculation of the power of a particular test proceeds as follows:

1. Determine the critical value,  $F_{k-1, N-k, \alpha}$  where  $\alpha$  is the probability of a type-I error and  $k$  and  $N$  are defined above. Note that this is a two-tailed test as no direction is assigned in the alternative hypothesis.
2. Choose an effect size parameter and the value of it that you want the experiment to be able to detect.
3. Compute the power as the probability of being greater than  $F_{k-1, N-k, \alpha}$  on a noncentral- $F$  distribution with noncentrality parameter  $\lambda$ , where  $\lambda$  is calculated from  $f$  or  $\eta^2$  using the formulas given above.

---

## Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

---

### Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

---

#### Solve For

##### Solve For

This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are *Effect Size*, *Sample Size*, *Alpha*, and *Power*. Under most situations, you will select either *Power* for a power analysis or *Sample Size* for sample size determination.

---

#### Power and Alpha

##### Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of equal means when in fact the means are different.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

##### Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected. In this procedure, a type-I error occurs when you reject the null hypothesis of equal means when in fact the means are equal.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

## One-Way Analysis of Variance F-Tests using Effect Size

### Sample Size / Groups

#### G (Number of Groups)

This is the number of group means being compared. It must be greater than or equal to two.

#### Group Allocation Ratios

Enter a list of positive, numeric values, one for each group. The sample size of group  $i$  is found by multiplying the  $i^{\text{th}}$  number from this list times the value of  $n$  and rounding up to the next whole number. The number of values must match the number of groups,  $G$ . When too few numbers are entered, 1's are added. When too many numbers are entered, the extras are ignored.

- **Equal**

If all sample sizes are to be equal, enter "Equal" here and the desired sample size in  $n$ . A set of  $G$  1's will be used. This will result in  $n_1 = n_2 = n_3 = n$ . That is, all sample sizes are equal to  $n$ .

#### n (Subjects per Group)

This is the base, per group, sample size. One or more values, separated by blanks or commas, may be entered. A separate analysis is performed for each value listed here.

The group sample sizes are determined by multiplying this number by each of the Group Sample Size Pattern numbers. If the Group Sample Size Pattern numbers are represented by  $m_1, m_2, m_3, \dots, m_G$  and this value is represented by  $n$ , the group sample sizes  $n_1, n_2, n_3, \dots, n_G$  are calculated as follows:

$$n_1 = [n(m_1)]$$

$$n_2 = [n(m_2)]$$

$$n_3 = [n(m_3)]$$

etc.

where the operator,  $[X]$  means the next integer after  $X$ , e.g.  $[3.1]=4$ .

For example, suppose there are three groups and the Group Sample Size Pattern is set to  $1,2,3$ . If  $n$  is 5, the resulting sample sizes will be 5, 10, and 15. If  $n$  is 50, the resulting group sample sizes will be 50, 100, and 150. If  $n$  is set to  $2,4,6,8,10$ , five sets of group sample sizes will be generated and an analysis run for each. These sets are:

2	4	6
4	8	12
6	12	18
8	16	24
10	20	30

#### Fractional Allocation Ratios

As a second example, suppose there are three groups and the Group Sample Size Pattern is  $0.2,0.3,0.5$ . When the fractional Pattern values sum to one,  $n$  can be interpreted as the total sample size of all groups ( $N$ ) and the allocation ratios as the proportion of the total in each group.

If  $n$  is 10, the three group sample sizes would be 2, 3, and 5.

If  $n$  is 20, the three group sample sizes would be 4, 6, and 10.

If  $n$  is 12, the three group sample sizes would be

$(0.2)12 = 2.4$  which is rounded up to the next whole integer, 3.

$(0.3)12 = 3.6$  which is rounded up to the next whole integer, 4.

$(0.5)12 = 6$ .

Note that in this case,  $3+4+6$  does not equal  $n$  (which is 12). This can happen because of rounding.

## One-Way Analysis of Variance F-Tests using Effect Size

---

**Effect Size****Effect Size Type**

Select the parameter to be used to specify the effect size. The possible choices are  $f$  and  $\eta^2$ .

 **$f$  ( $\sigma_m / \sigma$ )**

$f$  is a common measure of effect size as referenced in Cohen (1988). It is the ratio of the variation among the group means to the average variation among subjects within each group as measured by their standard deviations. The formula for calculating  $f$  was given above.

The values of  $f$  can range from zero upwards. Cohen (1988, 285-287) gives the following interpretation of  $f$  values. An  $f = 0.10$  is a *small* effect. An  $f = 0.25$  is a *medium* effect. An  $f = 0.40$  is a *large* effect.

You can enter a single value such as *0.1* or a series of values such as *0.1 0.25 0.4* or *0.1 to 0.4 by 0.1*. When a series of values is entered, **PASS** will generate a separate calculation result for each value of the series.

 **$\eta^2$  ( $\sigma_m^2 / (\sigma_m^2 + \sigma^2)$ )**

$\eta^2$  (or eta-squared) is an alternative measure of effect size that was defined above. The values of  $\eta^2$  can range from zero to one.

Using the relationship between  $f$  and  $\eta^2$ , we can defined small, medium, and large values of  $\eta^2$  as follows.

An  $\eta^2 = 0.0099 \approx 0.01$  is a *small* effect. An  $\eta^2 = 0.0588 \approx 0.06$  is a *medium* effect. And, an  $\eta^2 = 0.1379 \approx 0.14$  is a *large* effect.

You can enter a single value such as *0.1* or a series of values such as *0.01 0.06 0.14*. When a series of values is entered, **PASS** will generate a separate calculation result for each value of the series.

## One-Way Analysis of Variance F-Tests using Effect Size

## Example 1 – Finding Sample Size

An experiment is being designed to compare the means of four groups using an  $F$  test with a significance level of 0.05. The researchers would like to determine the sample size required to detect a small, medium, or large effect when the power is either 0.80 or 0.90.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **One-Way Analysis of Variance F-Tests using Effect Size** procedure. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Power.....	<b>0.8 0.9</b>
Alpha.....	<b>0.05</b>
G (Number of Groups).....	<b>4</b>
Group Allocation Ratios.....	<b>Equal</b>
Effect Size Type.....	<b>f = <math>\sigma_m/\sigma</math></b>
f ( $\sigma_m / \sigma$ ).....	<b>0.1 0.25 0.4</b>

### Output

Click the Calculate button to perform the calculations and generate the following output.

### Numeric Results

#### Numeric Results

Power	Average n	G	Total N	f	$\eta^2$	Alpha
0.8007	274.00	4	1096	0.1000	0.0099	0.0500
0.8040	45.00	4	180	0.2500	0.0588	0.0500
0.8234	19.00	4	76	0.4000	0.1379	0.0500
0.9007	356.00	4	1424	0.1000	0.0099	0.0500
0.9018	58.00	4	232	0.2500	0.0588	0.0500
0.9115	24.00	4	96	0.4000	0.1379	0.0500

#### References

- Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates. Hillsdale, New Jersey
- Desu, M. M. and Raghavarao, D. 1990. *Sample Size Methodology*. Academic Press. New York.
- Fleiss, Joseph L. 1986. *The Design and Analysis of Clinical Experiments*. John Wiley & Sons. New York.
- Kirk, Roger E. 1982. *Experimental Design: Procedures for the Behavioral Sciences*. Brooks/Cole. Pacific Grove, California.

#### Report Definitions

- Power is the probability of rejecting a false null hypothesis. It should be close to one.
- n is the average group sample size.
- G is the number of groups.
- Total N is the total sample size of all groups combined.
- f is the effect size measure calculated by  $\sigma_m / \sigma$ . Cohen (1988, 285-287) suggests Small = 0.1, Medium = 0.25, and Large = 0.4.
- $\eta^2$  is the effect size measure calculated by  $\sigma_m^2 / (\sigma_m^2 + \sigma^2)$ . It is the proportion of the variation in Y explained by the groups.
- Alpha is the probability of rejecting a true null hypothesis. It should be small.

## One-Way Analysis of Variance F-Tests using Effect Size

### Summary Statements

In a one-way ANOVA study, sample sizes of 274, 274, 274, and 274 are obtained from the 4 groups whose means are to be compared. The total sample of 1096 subjects achieves 80% power to detect differences among the means versus the alternative of equal means using an F test with a 0.0500 significance level. The size of the variation in the means is represented by the effect size  $f = \sigma_m / \sigma$ , which is 0.1000.

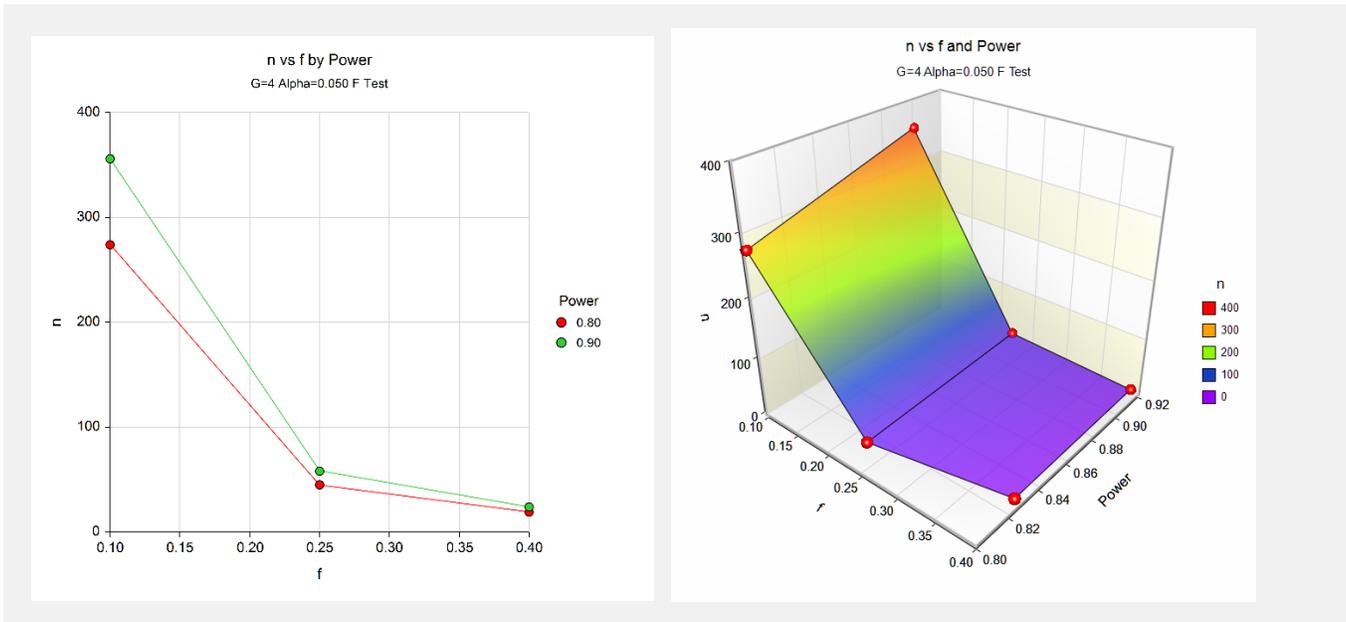
**Details when Alpha = 0.0500, Power = 0.8007, f = 0.1000,**

Group	Ni	Percent
		Ni of Total Ni
1	274.00	25.00
2	274.00	25.00
3	274.00	25.00
4	274.00	25.00
ALL	1096	100.00

.  
.
   
.

These reports show the sample size required to detect each effect size.

### Chart Section



These plots give a visual presentation to the results in the Numeric Report.

## One-Way Analysis of Variance F-Tests using Effect Size

**Example 2 – Validation using Fleiss (1986)**

Fleiss (1986) page 374 presents an example of determining a sample size in an experiment with 4 groups; means of 9.775, 12, 12, and 14.225; standard deviation of 3; alpha of 0.05, and beta of 0.20. He determines that a sample size of 11 per group is needed. These means and standard deviation result in a value of 0.5244 for  $f$ .

**Setup**

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **One-Way Analysis of Variance F-Tests using Effect Size** procedure. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Alpha.....	<b>0.05</b>
Power.....	<b>0.8</b>
G (Number of Groups).....	<b>4</b>
Group Allocation Ratios.....	<b>Equal</b>
Effect Size Type.....	<b><math>f = \sigma_m/\sigma</math></b>
$f$ ( $\sigma_m / \sigma$ ).....	<b>0.1 0.25 0.4</b>

**Output**

Click the Calculate button to perform the calculations and generate the following output.

**Numeric Results**

Numeric Results						
Power	Average n	G	Total N	f	$\eta^2$	Alpha
0.8027	11.00	4	44	0.5244	0.2157	0.0500

**PASS** also found  $n = 11$ . Note that Fleiss used calculations based on a normal approximation, but **PASS** uses exact calculations based on the non-central  $F$  distribution.