

Chapter 546

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 from the rest.

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.

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 within the groups 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 G groups each have a normal distribution and equal means ($\mu_1 = \mu_2 = \dots = \mu_G$). Let $N_1 = N_2 = \dots = N_G$ 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

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

Let σ denote the 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 also depends on the noncentrality parameter, λ . This parameter is calculated as

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

where

$$\sigma_m^2 = \sum_{i=1}^G \left(\frac{N_i}{N} \right) (\mu_i - \bar{\mu}_w)^2$$

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

$$\lambda = Nf^2$$

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 η^2 (*eta-squared*). This parameter is defined as

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

η^2 is the proportion of the total variation in the dependent variable that is attributable to the groups. The above equation can be solved for f as follows

$$f^2 = \frac{\eta^2}{1 - \eta^2}$$

Power Calculations for One-Way ANOVA

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

1. Determine the critical value, $F_{G-1, N-G, \alpha}$ where α is the probability of a type-I error and G 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 its value.
3. Compute the power as the probability of being greater than $F_{G-1, N-G, \alpha}$ on a noncentral- F distribution with noncentrality parameter λ , where λ is calculated from f or η^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 Power, Sample Size, and Effect Size. Under most situations, you will select either *Power* for a power analysis or *Sample Size* for a 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 also 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*.

Sample Size and Group Allocation

G (Number of Groups)

This is the number of groups (arms) whose means are being compared. The number of items used in the Group Allocation boxes is controlled by this number.

This value must be an integer greater than or equal to two.

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

Group Allocation Input Type (when Solve For = Power or Effect Size)

Specify how you want to enter the information about how the subjects are allocated to each of the G groups.

Possible options are:

- **Equal ($N_1 = N_2 = \dots = N_G$)**
The sample size of all groups is N_i . Enter one or more values for the common group sample size.
- **Enter group multipliers**
Enter a list of group multipliers (r_1, r_2, \dots, r_G) and one or more values of N_i . The individual group sample sizes are found by multiplying the multipliers by N_i . For example, $N_1 = r_1 \times N_i$.
- **Enter N_1, N_2, \dots, N_G**
Enter a list of group sample sizes, one for each group.
- **Enter columns of N_i 's**
Select one or more columns of the spreadsheet that each contain a set of group sample sizes going down the column. Each column is analyzed separately.

N_i (Subjects Per Group)

Enter N_i , the number of subjects in each group. The total sample size, N , is equal to $N_i \times G$.

You can specify a single value or a list.

Single Value

Enter a value for the individual sample size of all groups. If you enter '10' here and there are five groups, then each group will be assigned 10 subjects and the total sample size will be 50.

List of Values

A separate power analysis is calculated for each value of N_i in the list. All analyses assume that the common, group sample size is N_i .

Range of N_i

$N_i > 1$

Group Multipliers (r_1, r_2, \dots, r_G)

Enter a set of G multipliers, one for each group.

The individual group sample sizes is computed as $N_g = \text{ceiling}[r_g \times N_i]$, where $\text{ceiling}[y]$ is the first integer greater than or equal to y . For example, the multipliers $\{1, 1, 2, 2.95\}$ and *base* N_i of 10 would result in the sample sizes $\{10, 10, 20, 30\}$.

Incomplete List

If the number of items in the list is less than G , the missing multipliers are set equal to the last entry in the list.

Range

The items in the list must be positive. The resulting sample sizes must be at least 1.

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

Ni (Base Subjects Per Group)

Enter N_i , the base sample size of each group. The number of subjects in the group is found by multiplying this number by the corresponding group multiplier, $\{r_1, r_2, \dots, r_G\}$, and rounding up to the next integer.

You can specify a single value or a list.

Single Value

Enter a value for the base group subject count.

List of Values

A separate power analysis is calculated for each value of N_i in the list.

Range

$\text{Ceiling}[N_i \times r_i] \geq 1$.

N1, N2, ..., NG (List)

Enter a list of G subject counts, one for each group.

Incomplete List

If the number of items in the list is less than G, the missing subject counts are set equal to the last entry in the list.

Range

The items in the list must be positive. At least one item in the list must be greater than 1.

Columns of Ni's

Enter one or more spreadsheet columns containing vertical lists of group subject counts.

Press the Spreadsheet icon (directly to the right) to select the columns and then enter the values.

Press the Input Spreadsheet icon (to the right and slightly up) to view/edit the spreadsheet. Also note that you can obtain the spreadsheet by selecting "Tools", then "Input Spreadsheet", from the menus.

On the spreadsheet, the group subject counts are entered going down.

Examples (assuming G = 3)

C1	C2	C3
111	1	28
115	20	68
100	30	46

Definition of a Single Column

Each column gives one list. Each column results in a new scenario. The columns are not connected, but all should have exactly G rows.

Each entry in the list is the subject count of that group.

Incomplete List

If the number of items in the list is less than G, the missing entries are set equal to the last entry in the list.

Valid Entries

All values should be positive integers. At least one value must be greater than one.

Note

The column names (C1, C2, ...) can be changed by right-clicking on them in the spreadsheet.

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

Group Allocation Input Type (when Solve For = Sample Size)

Specify how you want to enter the information about how the subjects are allocated to each of the G groups.

Options

- **Equal ($N_1 = N_2 = \dots = N_G$)**
All group subject counts are equal to N_i . The value of N_i will be found by conducting a search.
- **Enter group allocation pattern**
Enter an allocation pattern (r_1, r_2, \dots, r_G). The pattern consists of a set of G numbers. These numbers will be rescaled into proportions by dividing each item by the sum of all items. The individual group subject counts are found by multiplying these proportions by N (the total subject count) and rounding up.
- **Enter columns of allocation patterns**
Select one or more columns of the spreadsheet that each contain a group allocation pattern going down the column. Each column is analyzed separately.

Group Allocation Pattern (r_1, r_2, \dots, r_G)

Enter an allocation pattern (r_1, r_2, \dots, r_G). The pattern consists of a set of G numbers. These numbers will be rescaled into proportions of N by dividing each item by the sum of all items. The individual group subject counts are found by multiplying these proportions by N (the total subject count) and rounding up.

For example, the pattern $\{1, 3, 4\}$ will be rescaled to $\{0.125, 0.375, 0.5\}$. The group subject counts will be constrained to these proportions (within rounding) during the search for the subject count configuration that meets the power requirement.

Incomplete List

If the number of items in the list is less than G , the missing numbers are set equal to the last entry in the list.

Range

The items in the list must be positive. The resulting subject counts must be at least 1.

Columns of Group Allocation Patterns

Enter one or more spreadsheet columns containing vertical lists of group allocation patterns.

Press the Spreadsheet icon (directly to the right) to select the columns and then enter the values.

Press the Input Spreadsheet icon (to the right and slightly up) to view/edit the spreadsheet. Also note that you can obtain the spreadsheet by selecting "Tools", then "Input Spreadsheet", from the menus.

On the spreadsheet, the group allocation patterns are entered going down.

Examples (assuming $G = 3$)

C1	C2	C3
1	1	3
1	2	3
3	2	1

Definition of a Single Column

Each column gives one allocation pattern. Each column results in a new scenario. The columns are not connected, but all should have exactly G rows.

Incomplete List

If the number of items in a list is less than G , the missing numbers are set equal to the last entry in the list before they are rescaled.

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

Valid Entries

All values should be positive numbers. You can enter decimal values.

Note

The column names (C1, C2, ...) can be changed by right-clicking on them in the spreadsheet.

Effect Size

Effect Size Type

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

f (σ_m / σ)

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.

η^2 ($\sigma_m^2 / (\sigma_m^2 + \sigma^2)$)

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

Using the relationship between f and η^2 , we can define small, medium, and large values of η^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.

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 procedure window. 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>
Design Tab	
Solve For	Sample Size
Power.....	0.8 0.9
Alpha.....	0.05
G (Number of Groups).....	4
Group Allocation Input Type	Equal (N1 = N2 = ... = NG)
Effect Size Type.....	f = σ_m/σ
f (σ_m / σ).....	0.1 0.25 0.4

Annotated Output

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

Numeric Results

Numeric Results

Number of Groups: 4

	Total Sample Size N	Group Alloc Set ri Set	Effect Size f	Alpha
Power				
0.8007	1096	ri(1)	0.1000	0.050
0.8040	180	ri(1)	0.2500	0.050
0.8234	76	ri(1)	0.4000	0.050
0.9007	1424	ri(1)	0.1000	0.050
0.9018	232	ri(1)	0.2500	0.050
0.9115	96	ri(1)	0.4000	0.050

Set(Set Number): Values

ri(1): 0.250, 0.250, 0.250, 0.250

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.
- Total Sample Size N is the total number of subjects in the study.
- Group Alloc Set ri gives the name and number of the set containing the allocation proportion of each group.
- Effect Size f is a measure of the effect size. It is the ratio of σ_m and σ .
- Alpha is the significance level of the test: the probability of rejecting the null hypothesis of equal means when it is true.

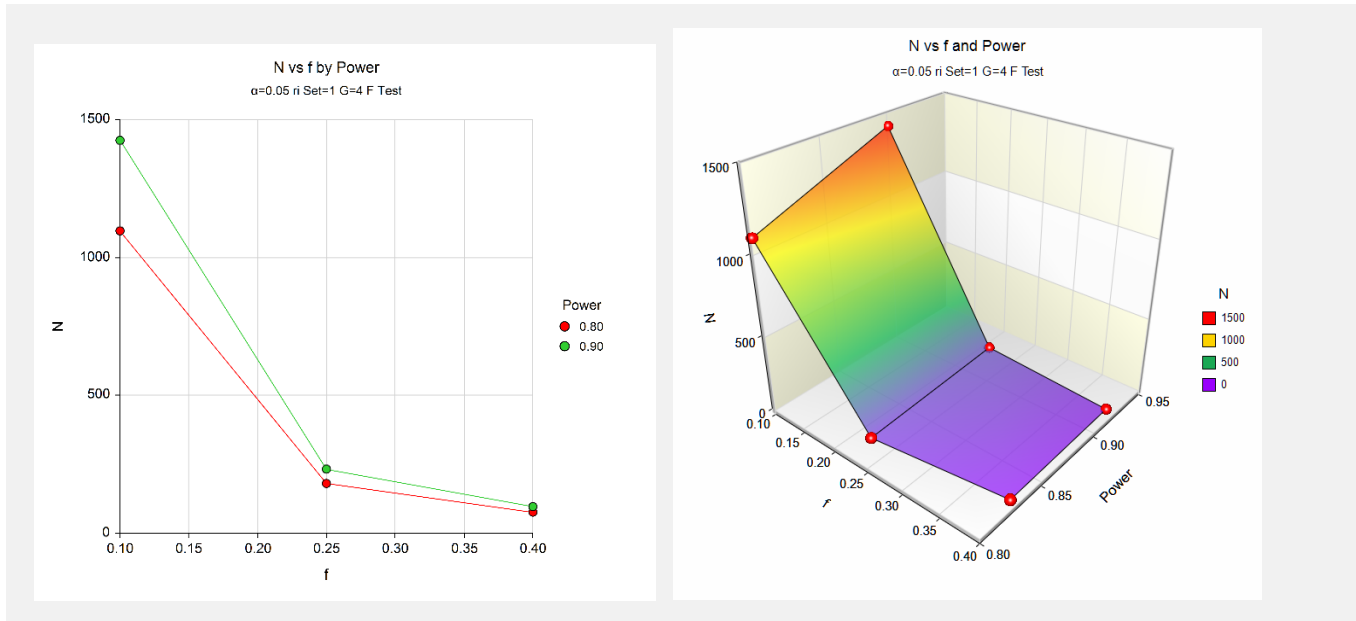
One-Way Analysis of Variance F-Tests using Effect Size

Summary Statements

In a one-way ANOVA study, a sample of 1096 subjects, divided among 4 groups, achieves a power of 0.8007. This power assumes an F test is used with a significance level of 0.050. The group subject proportions are 0.250, 0.250, 0.250. The effect size f , which is calculated using $f = (\sigma_m / \sigma)$, is equal to 0.1000.

This report shows the numeric results of this power study.

Plots Section



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

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—for a total sample size of 44. 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 procedure window. 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>
Design Tab	
Solve For	Sample Size
Power.....	0.8
Alpha.....	0.05
G (Number of Groups).....	4
Group Allocation Input Type	Equal (N1 = N2 = ... = NG)
Effect Size Type.....	$f = \sigma_m / \sigma$
$f (\sigma_m / \sigma)$	0.5244

Output

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

Numeric Results

Numeric Results				
Number of Groups: 4				
	Total Sample Size	Group Alloc Set	Effect Size	
Power	N	ri Set	f	Alpha
0.8027	44	ri(1)	0.5244	0.050
Set(Set Number): Values				
ri(1): 0.250, 0.250, 0.250, 0.250				

PASS also found $N = 44$. Note that Fleiss used calculations based on a normal approximation, but PASS uses exact calculations based on the non-central F distribution.