Chapter 587

Non-Zero Null Tests for One-Way Analysis of Variance Assuming Equal Variances

Introduction

This procedure computes power and sample size of non-zero null F-tests of the means of two or more groups which are analyzed using a noncentral F-test. The results in this chapter come from Shieh (2018), Jan and Shieh (2019), and Cohen (1988).

Background

The common way of comparing the means of two or more groups sampled in a one-way design is with an F-test. One complaint about this test is that no matter how slight the differences among the means, since the null hypothesis is that there is absolutely no difference among the means, it is always possible to find statistical significance with a large-enough sample size. One way to combat this problem is to adjust the hypothesis to ignore the possibility that small, negligible effects are found to be significant. This is accomplished using an F-test with a non-zero null hypothesis. Another way of saying this is that the test finds only minimally important mean differences.

Technical Details for the One-Way ANOVA Test

Suppose *G* groups each have a normal distribution and with means $\mu_1, \mu_2, \ldots, \mu_G$ and common variance σ^2. Let $n_1, n_2, \ldots, n_G$ denote the sample size of each group and let *N* denote the total sample size of all groups. The non-zero null F-test requires one to show that the means are sufficiently different from each other. Shieh (2018) accomplished this by defining a set of non-zero null means, $\mu_{01}, \mu_{02}, \ldots, \mu_{0G}$, that are as close together as possible and still be *different*.

He then summarized these means using their weighted variance

$$\sigma_{m0}^2 = \sum_{i=1}^{G} \left(\frac{n_i}{N}\right) (\mu_{0i} - \bar{\mu}_0)^2$$

where $\bar{\mu}_0$ is the weighted mean

$$\bar{\mu}_0 = \sum_{i=1}^{G} \left(\frac{n_i}{N}\right) \mu_{0i}$$

The corresponding variance of the group means under the alternative hypothesis is given by

$$\sigma_{m1}^2 = \sum_{i=1}^{G} \left(\frac{n_i}{N}\right)(\mu_{1i} - \bar{\mu}_1)^2$$

where $\bar{\mu}_1$ is the weighted mean

$$\bar{\mu}_1 = \sum_{i=1}^{G} \left(\frac{n_i}{N}\right)\mu_{1i}$$

Given the above terminology, Shieh (2018) suggested testing the hypothesis of minimally important mean difference using

$$H_0: f \leq f_0 \quad \text{versus} \quad H_1: f > f_0$$

where $f^2 = \sigma_m^2/\sigma^2$ represents the usual F-ratio from a one-way design and $f_0^2 = \sigma_{m0}^2/\sigma^2$ is the non-zero null bound defined above.

Under H_0, the usual F statistic, denoted F^*, is assumed to follow the noncentral-F distribution

$$F^* \sim F'_{G-1,N-G,\Lambda_0}$$

where $\Lambda_0 = Nf_0^2$. Note that Λ_0 does not depend on data. Rather, it depends on the user specified set of non-zero null values $\mu_{01}, \mu_{02}, \dots, \mu_{0G}$.

The value of F^* is given by

$$F^* = \frac{SSM/(G-1)}{SSE/(N-G)}$$

where *SSM* is the sum of squares of treatment means and *SSE* is the sum of squares of error.

The null hypothesis is rejected at the significance level α if $F^* > F'_{1-\alpha,G-1,N-G,\Lambda_0}$.

The power function of this test statistic at a particular set of means is given by

$$\text{Power} = \Pr\left[F'_{G-1,N-G,\Lambda_1} > F'_{1-\alpha,G-1,N-G,\Lambda_0}\right]$$

where $\Lambda_1 = Nf_1^2$ and $f_1^2 = \sigma_{m1}^2/\sigma^2$. This can easily be computed using a noncentral-*F* cumulative distribution function. You have to be careful that $\sigma_{m0}^2 < \sigma_{m1}^2$.

If a sample size is desired, it can be determined using a standard binary search algorithm.

Specifying σ_m^2

The statistical hypotheses involve the variance of the means, σ_m^2. This quantity may be hard to interpret and thus difficult to come up with during the study planning phase. Cohen (1988) suggested using a different parameter that is easier to interpret. This parameter is the *range* given by

$$\delta = \mathrm{Max}\{\mu_1, \mu_2, \dots, \mu_G\} - \mathrm{Min}\{\mu_1, \mu_2, \dots, \mu_G\}$$

Note that the range of a set of means easier to understand than their variance or standard deviation. For example, suppose a set of means has the values {2, 3, 2, 1}. Most people could quickly grasp the meaning of the fact that the range of these means is 2 but would find it more difficult to understand that the standard deviation is 0.7071. So, even though the hypotheses are stated in terms of the standard deviations of the means, at the planning stage, it is much easier to agree on target values for the ranges under the null and alternative hypotheses.

Unfortunately, except for the case when G = 2, there is no specific relationship between δ and σ_m^2. Upon closer investigation, it has been found that there is a relationship between σ_m^2 and the range of possible values of δ. Specifically, when all group sample sizes are identical, Cohen (1988, chapter 8) gives this relationship

1. The minimum σ_m for a particular value of δ occurs when means of equal magnitude are placed at each end point and the remaining means are set to their average. This pattern of means, apart from a constant, is $\left\{\frac{-\delta}{2}, 0, \dots, 0, \frac{\delta}{2}\right\}$. In this case, $\sigma_{Min} = \delta/\sqrt{2G}$.
2. The maximum σ_m occurs when two means whose difference is δ are placed at the end points. When G is even, this pattern of means, apart from a constant, is $\left\{\frac{-\delta}{2}, \dots, \frac{-\delta}{2}, \frac{\delta}{2}, \dots, \frac{\delta}{2}\right\}$. This configuration gives $\sigma_{Max} = \delta/2$ if G is even and $\sigma_{Max} = \delta\sqrt{G^2-1}/2G$ if G is odd.

Examples of the Relationship between δ and σ_m

G	δ	Min Means	σ_{Min}	Max Means	σ_{Max}
4	2	{-1,0,0,1}	0.7071	{-1,-1,1,1}	1.0
5	2	{-1,0,0,0,1}	0.6325	{-1,-1,1,1,1}	0.9798
6	2	{-1,0,0,0,0,1}	0.5774	{-1,-1,-1,1,1,1}	1.0
7	2	{-1,0,0,0,0,0,1}	0.5345	{-1,-1,-1,1,1,1,1}	0.9897

Specifying a Reasonable Effect Size

Here are a reasonable set of steps to take when specifying a valid effect size for this procedure.

1. Find a reasonable value of σ. This gives you the scale.
2. Select a value for the minimally important range. This is the maximum range that you are willing to interpret as being negligible.
3. Find the minimum and maximum possible values of σ_m for the range in step 2.
4. Evaluate whether the mean pattern is closer to the *Min Means* or *Max Means* and then plan accordingly.
5. Use this information to set a range of possible σ_{m0} values.
6. Go through similar steps to determine an appropriate set of values for σ_{m1}.

Example 1 – Finding Power

An experiment is being designed to assess the equality of the means of four groups using an *F* test at a significance level of 0.05. Previous studies have shown that the standard deviation is about 2.

The researchers decide to set the minimally important difference to 1. Using the Data tab of Standard Deviation Estimator tool, they find the minimum possible value of $\sigma m0$ for four values with a range of 1 is 0.3536 and the maximum is 0.5. For this run, the researchers set this value to 0.43 which is near the middle of the possible values.

The researchers decide to set the minimally important difference to be used for the alternative hypothesis to 2. Using the Data tab of Standard Deviation Estimator tool, they find the minimum possible value of $\sigma m1$ for four values with a range of 2 is 0.7071 and the maximum is 1. The researchers set this option to {0.7, 0.8, 0.9, 1}

The researchers want to compute the power for several group sample sizes between 20 and 80. These sample sizes will be equal across all groups.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For .... **Power**
Alpha.. **0.05**
G (Number of Groups) **4**
Group Allocation Input Type .... **Equal to ni (Sample Size per Group)**
ni (Sample Size per Group) **20 40 60 80**
μ0 Input Type.. **Enter σm0 (SD of μ0)**
σm0 (SD of μ0) **0.43**
μ1 Input Type.. **Enter σm1 (SD of μ1)**
σm1 (SD of μ1) **0.7 0.8 0.9 1**
σ (Standard Deviation).................................. **2**

Output

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

Numeric Reports

Numeric Results

Solve For: Power
Number of Groups: 4

	Sample Size		SD of Means			
Power	Total N	Group ni	H0 σm0	H1 σm1	Standard Deviation σ	Alpha
0.28351	80	20	0.43	0.7	2	0.05
0.43402	80	20	0.43	0.8	2	0.05
0.59599	80	20	0.43	0.9	2	0.05
0.74351	80	20	0.43	1.0	2	0.05
0.47823	160	40	0.43	0.7	2	0.05
0.70596	160	40	0.43	0.8	2	0.05
0.87286	160	40	0.43	0.9	2	0.05
0.95908	160	40	0.43	1.0	2	0.05
0.63163	240	60	0.43	0.7	2	0.05
0.85906	240	60	0.43	0.8	2	0.05
0.96546	240	60	0.43	0.9	2	0.05
0.99478	240	60	0.43	1.0	2	0.05
0.74639	320	80	0.43	0.7	2	0.05
0.93619	320	80	0.43	0.8	2	0.05
0.99147	320	80	0.43	0.9	2	0.05
0.99942	320	80	0.43	1.0	2	0.05

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N The total number of subjects in the study.
ni The Sample Size per Group is the number of items sampled from each group in the study.
σm0 The SD of Group Means|H0 is the standard deviation of the group means assumed by the null hypothesis. Note that this value also depends on the group sample sizes.
σm1 The SD of Group Means|H1 is the standard deviation of the group means assumed by the alternative hypothesis. Note that this value also depends on the group sample sizes.
σ The common standard deviation of the responses within a group.
Alpha The probability of rejecting a true null hypothesis.

Group Sample Size Details

n	N	Group Sample Sizes	Group Allocation Proportions
n(1)	80	20, 20, 20, 20	0.25, 0.25, 0.25, 0.25
n(2)	160	40, 40, 40, 40	0.25, 0.25, 0.25, 0.25
n(3)	240	60, 60, 60, 60	0.25, 0.25, 0.25, 0.25
n(4)	320	80, 80, 80, 80	0.25, 0.25, 0.25, 0.25

Summary Statements

In a one-way ANOVA study with a non-zero null boundary, a sample of 80 subjects, divided among 4 groups, achieves a power of 28%. This power assumes a non-central F test with a significance level of 0.05. The group subject counts are 20, 20, 20, 20. The standard deviation of hypothesized means under the null hypothesis is 0.43. The standard deviation of hypothesized means under the alternative hypothesis is 0.7. The common standard deviation of the responses is 2.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	80	100	20
20%	160	200	40
20%	240	300	60
20%	320	400	80

Dropout Rate The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
N The evaluable sample size at which power is computed (as entered by the user). If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated power.
N' The total number of subjects that should be enrolled in the study in order to obtain N evaluable subjects, based on the assumed dropout rate. N' is calculated by inflating N using the formula N' = N / (1 - DR), with N' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
D The expected number of dropouts. D = N' - N.

Dropout Summary Statements

Anticipating a 20% dropout rate, 100 subjects should be enrolled to obtain a final sample size of 80 subjects.

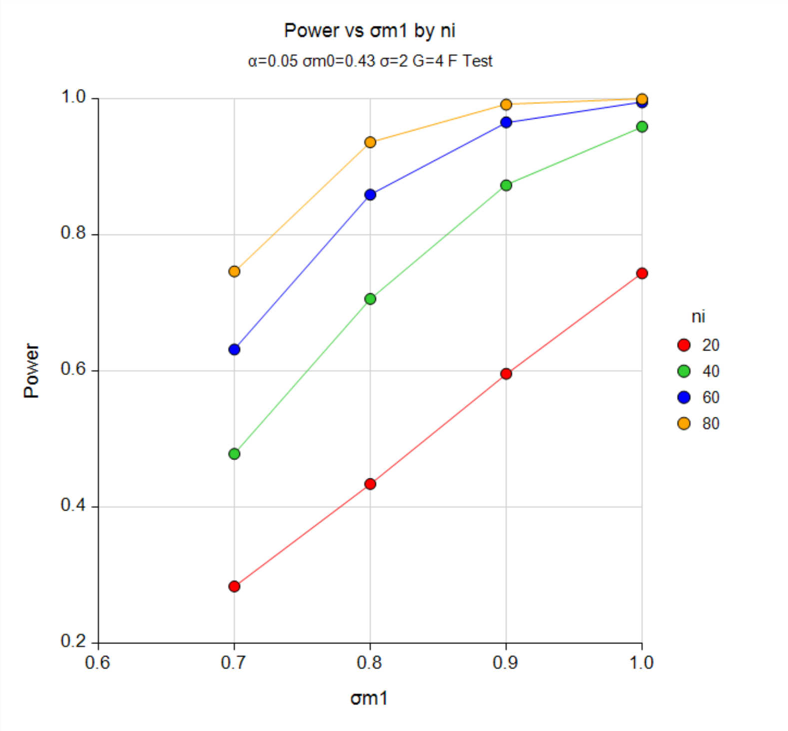
References

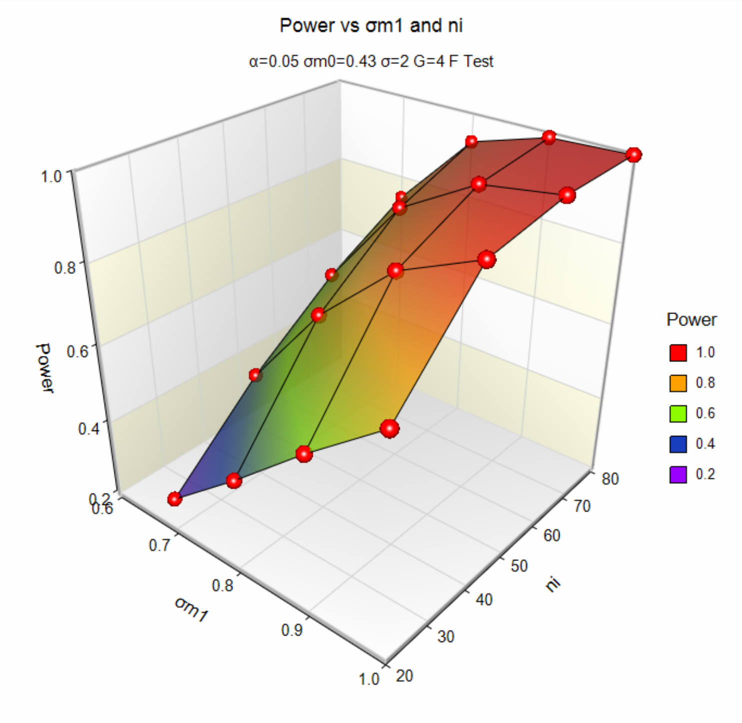
Shieh, G. 2018. 'On Detecting a Minimal Important Difference among Standardized Means'. Current Psychology, Vol 37, Pages 640-647. Doi: 10.1007/s12144-016-9549-5
Jan, S-L and Shieh, G. 2019. 'On the Extended Welch Test for Assessing Equivalence of Standardized Means'. Statistics in Biopharmaceutical Research. DOI:10.1080/19466315.2019.1654915
Cohen, Jacob. 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates. Hillsdale, New Jersey.

This report shows the numeric results of this power study.

Plots Section

Plots





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

Example 2 – Finding the Sample Size Necessary to Reject

Continuing with the last example, we will determine how large the sample size would need to be to achieve a power of 0.9.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For .... **Sample Size**
Power.. **0.90**
Alpha... **0.05**
G (Number of Groups) **4**
Group Allocation Input Type .... **Equal (n1 = ··· = nG)**
μ0 Input Type.. **Enter σm0 (SD of μ0)**
σm0 (SD of μ0) **0.43**
μ1 Input Type.. **Enter σm1 (SD of μ1)**
σm1 (SD of μ1) **0.7 0.8 0.9 1**
σ (Standard Deviation).................................. **2**

Output

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

Numeric Results

Solve For: Sample Size
Number of Groups: 4

	Sample Size		SD of Means			
Power	**Total N**	**Group ni**	**H0 σm0**	**H1 σm1**	**Standard Deviation σ**	**Alpha**
0.90004	504	126	0.43	0.7	2	0.05
0.90073	276	69	0.43	0.8	2	0.05
0.90109	176	44	0.43	0.9	2	0.05
0.90310	124	31	0.43	1.0	2	0.05

Group Sample Size Details

n	N	Group Sample Sizes	Group Allocation Proportions
n(1)	504	126, 126, 126, 126	0.25, 0.25, 0.25, 0.25
n(2)	276	69, 69, 69, 69	0.25, 0.25, 0.25, 0.25
n(3)	176	44, 44, 44, 44	0.25, 0.25, 0.25, 0.25
n(4)	124	31, 31, 31, 31	0.25, 0.25, 0.25, 0.25

The required sample size varies from 124 to 504 as the value of σm1 is increased.

Example 3 – Validation using Shieh (2018)

Shieh (2018) page 644 presents an example of finding power when alpha = 0.05, G = 3, σ = 3.189, σm0 = 0.3189, and μ1 = {7.77 9.77 6.68}, and ni = 22. The resulting power is reported as 0.7109.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For	**Power**
Alpha	**0.05**
G (Number of Groups)	**3**
Group Allocation Input Type	**Equal to ni (Sample Size per Group)**
ni (Sample Size per Group)	**22**
μ0 Input Type	**Enter σm0 (SD of μ0)**
σm0 (SD of μ0)	**0.3189**
μ1 Input Type	**Enter μ1 (Group Means|H1)**
μ1 (Group Means|H1)	**7.77 9.77 6.68**
σ (Standard Deviation)	**3.189**

Output

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

Numeric Results

Solve For: Power
Number of Groups: 3

	Sample Size			SD of Means			
Power	Total N	Group ni	H1 Pattern μ1	H0 σm0	H1 σm1	Standard Deviation σ	Alpha
0.7109	66	22	μ1(1)	0.3189	1.27959	3.189	0.05

Item	Values
μ1(1)	7.77, 9.77, 6.68

PASS also found the power to be 0.7109. Thus, the procedure is validated.