Chapter 419

Two-Sample T-Tests using Effect Size

Introduction

This procedure provides sample size and power calculations for one- or two-sided two-sample t-tests when the effect size is specified rather than the means and variance(s). The details of procedure are given in Cohen (1988). The design corresponding to this test procedure is sometimes referred to as a *parallel-groups* design. In this design, two groups from independent, normally distributed populations are compared by considering the difference in their means scaled be their common standard deviation.

This procedure is specific to the two-sample t-test assuming equal variance. If the variances are known to be significantly different, this procedure can still be used if the group sample sizes are equal and the average of the variances is used.

Test Assumptions

When running a two-sample equal-variance t-test, the basic assumptions are that the distributions of the two populations are approximately normal, and that the variances of the two distributions are the same. If the variances are different, this procedure can still be used if the two group sample sizes are nearly equal.

Test Procedure

If we assume that μ_1 and μ_2 represent the means of the two populations of interest and their common (unknown) standard deviation is σ, the effect size is represented by d where

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

The null hypothesis is H_0: d = 0 and the alternative hypothesis depends on the number of "sides" of the test:

Two-Sided:	$H_1: d \neq 0$ or $H_1: \mu_1 - \mu_2 \neq 0$
Upper One-Sided:	$H_1: d > 0$ or $H_1: \mu_1 - \mu_2 > 0$
Lower One-Sided:	$H_1: d < 0$ or $H_1: \mu_1 - \mu_2 < 0$

A suitable Type I error probability (α) is chosen for the test, the data is collected, and a *t*-statistic is generated using the formula:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

This t-statistic follows a t distribution with $n_1 + n_2 - 2$ degrees of freedom. The null hypothesis is rejected in favor of the alternative if,

for $H_1: d \neq 0$ or $H_1: \mu_1 - \mu_2 \neq 0$

$$t < t_{\alpha/2} \text{ or } t > t_{1-\alpha/2},$$

for $H_1: d > 0$ or $H_1: \mu_1 - \mu_2 > 0$

$$t > t_{1-\alpha},$$

Or, for $H_1: d < 0$ or $H_1: \mu_1 - \mu_2 < 0$

$$t < t_{\alpha}.$$

Comparing the t-statistic to the cut-off t-value (as shown here) is equivalent to comparing the p-value to α.

Power Calculation

The power is calculated using the same formulation as in the *Two-Sample T-Tests Assuming Equal Variances* procedure with the modification that the σ used in that procedure is set equal to one.

If the variances cannot be assumed to be equal, the modification suggested by Cohen (1988) is used. This modification is to substitute an average value of the two variances and then proceeding as if the variances were equal. The average value is computed using

$$\sigma = \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}$$

Cohen remarks that this method is only accurate if the two sample sizes are (nearly) equal.

The Effect Size

If we assume that μ_1 and μ_2 represent the means of the two populations of interest and their common (unknown) standard deviation is σ, the effect size is represented by d where

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

Cohen (1988) proposed the following interpretation of the d values. A d near 0.2 is a *small* effect, a d near 0.5 is a *medium* effect, and a d near 0.8 is a *large* effect. These values for small, medium, and large effects are popular in the social sciences. However, this convention is not as popular among the medical sciences since the scale of the effect is left unstated which makes interpretation difficult.

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*, *Effect Size*, and *Alpha*. In most situations, you will likely select either *Power* or *Sample Size*.

The 'Solve For' parameter is the parameter that will be displayed on the vertical axis of any plots that are shown.

Test Direction

Alternative Hypothesis

Specify whether the alternative hypothesis of the test is one-sided or two-sided. If a one-sided test is chosen, the hypothesis test direction is chosen based on whether the effect size is greater than or less than zero.

Two-Sided Hypothesis Test
H0: $d = 0$ vs. H1: $d \neq 0$

One-Sided Hypothesis Tests
Upper: H0: $d \leq 0$ vs. H1: $d > 0$
Lower: H0: $d \geq 0$ vs. H1: $d < 0$

Power and Alpha

Power

Power is the probability of rejecting the null hypothesis when it is false. Power is equal to 1 - Beta, so specifying power implicitly specifies beta. Beta is the probability obtaining a false negative with the statistical test. That is, it is the probability of accepting a false null hypothesis.

The valid range is 0 to 1. Different disciplines have different standards for setting power. The most common choice is 0.90, but 0.80 is also popular.

You can enter a single value, such as 0.90, or a series of values, such as .70 .80 .90, or .70 to .90 by .1.

When a series of values is entered, PASS will generate a separate calculation result for each value of the series.

Alpha

Alpha is the probability of obtaining a false positive with the statistical test. That is, it is the probability of rejecting a true null hypothesis. The null hypothesis is usually that the parameters of interest (means, proportions, etc.) are equal.

Since Alpha is a probability, it is bounded by 0 and 1. Commonly, it is between 0.001 and 0.10. Alpha is often set to 0.05 for two-sided tests and to 0.025 for one-sided tests. You can enter a single value, such as 0.05, or a series of values, such as .05 .10 .15, or .05 to .15 by .01. When a series of values is entered, PASS will generate a separate calculation result for each value of the series.

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on *N1* or *N2* or both.

The options are

- **Equal (N1 = N2)**

  This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.

- **Enter N1, solve for N2**

  Select this option when you wish to fix *N1* at some value (or values), and then solve only for *N2*. Please note that for some values of *N1*, there may not be a value of *N2* that is large enough to obtain the desired power.

- **Enter N2, solve for N1**

  Select this option when you wish to fix *N2* at some value (or values), and then solve only for *N1*. Please note that for some values of *N2*, there may not be a value of *N1* that is large enough to obtain the desired power.

- **Enter R = N2/N1, solve for N1 and N2**

  For this choice, you set a value for the ratio of *N2* to *N1*, and then **PASS** determines the needed *N1* and *N2*, with this ratio, to obtain the desired power. An equivalent representation of the ratio, *R*, is

  $$N2 = R * N1.$$

- **Enter percentage in Group 1, solve for N1 and N2**

  For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed *N1* and *N2* with this percentage to obtain the desired power.

N1 (Sample Size, Group 1)

This option is displayed if Group Allocation = "Enter N1, solve for N2"

N1 is the number of items or individuals sampled from the Group 1 population.

N1 must be ≥ 2. You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed if Group Allocation = "Enter N2, solve for N1"

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter R = N2/N1, solve for N1 and N2."

R is the ratio of *N2* to *N1*. That is,

$$R = N2 / N1.$$

Use this value to fix the ratio of *N2* to *N1* while solving for *N1* and *N2*. Only sample size combinations with this ratio are considered.

N2 is related to *N1* by the formula:

$$N2 = [R \times N1],$$

where the value *[Y]* is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1 (e.g., $N1 = 10$ and $N2 = 20$, or $N1 = 50$ and $N2 = 100$).

R must be greater than 0. If $R < 1$, then $N2$ will be less than $N1$; if $R > 1$, then $N2$ will be greater than $N1$. You can enter a single or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter percentage in Group 1, solve for N1 and N2."

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for $N1$ and $N2$. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

Sample Size (When <u>Not</u> Solving for Sample Size)

Group Allocation

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal (N1 = N2)**

  This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

- **Enter N1 and N2 individually**

  This choice permits you to enter different values for *N1* and *N2*.

- **Enter N1 and R, where N2 = R * N1**

  Choose this option to specify a value (or values) for *N1*, and obtain *N2* as a ratio (multiple) of *N1*.

- **Enter total sample size and percentage in Group 1**

  Choose this option to specify a value (or values) for the total sample size (*N*), obtain *N1* as a percentage of *N*, and then *N2* as *N - N1*.

Sample Size Per Group

This option is displayed only if Group Allocation = "Equal (N1 = N2)."

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for *N1*, and also the value for *N2*.

The Sample Size Per Group must be ≥ 2. You can enter a single value or a series of values.

N1 (Sample Size, Group 1)

*This option is displayed if Group Allocation = "Enter N1 and N2 individually" or "Enter N1 and R, where N2 = R * N1."*

N1 is the number of items or individuals sampled from the Group 1 population.

N1 must be ≥ 2. You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed only if Group Allocation = "Enter N1 and N2 individually."

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

*This option is displayed only if Group Allocation = "Enter N1 and R, where N2 = R * N1."*

R is the ratio of *N2* to *N1*. That is,

$$R = N2/N1$$

Use this value to obtain *N2* as a multiple (or proportion) of *N1*.

N2 is calculated from *N1* using the formula:

$$N2=[R \times N1],$$

where the value *[Y]* is the next integer ≥ *Y*.

For example, setting *R* = 2.0 results in a Group 2 sample size that is double the sample size in Group 1.

R must be greater than 0. If *R* < 1, then *N2* will be less than N1; if *R* > 1, then *N2* will be greater than *N1*. You can enter a single value or a series of values.

Total Sample Size (N)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines *N1* and *N2*.

The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.

You can enter a single value or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Effect Size

d

Enter one or more values for d, the effect size, that you wish to detect. This is a standardized difference between the group means. The effect size is calculated using

$d = (μ1 - μ2) / σ$

where μ1 and μ2 are the group means assumed by the alternative hypothesis and σ is your estimate of the population standard deviation.

The value of d can be any non-zero value (positive or negative). However, it is usually between -3 and 3, excluding 0.

You can enter a single value such as 0.5 or a series of values such as 0.2 0.5 0.8 or 0.2 to 0.8 by 0.1. When a series of values is entered, PASS will generate a separate calculation result for each value of the series.

Two-Sample T-Tests using Effect Size

Cohen's Effect Size Table

Cohen (1988) gave the following interpretation of d values that is still popular.

- **Small**

  $d = 0.2$ or 20% of σ

- **Medium**

  $d = 0.5$ or 50% of σ

- **Large**

  $d = 0.8$ or 80% of σ

Unequal Group Variances

If the two group variance are markedly different, the average of the two variances is computed and used:

$$\sigma = \sqrt{[(\sigma_1^2 + \sigma_2^2)/2]}$$

This estimate of σ requires that the two sample sizes, N1 and N2, be (nearly) equal. If they are not, this method may be inaccurate.

Using d is criticized because it obscures the true scale of the data.

Example 1 – Finding the Sample Size

Researchers wish to compare two types of local anesthesia using a balanced, parallel-group design. Subjects in pain will be randomized to one of two treatment groups, the treatment will be administered, and the subject's evaluation of pain intensity will be measured on a seven-point scale.

The researchers would like to determine the sample sizes required to detect a small, medium, and large effect size with a two-sided t-test when the power is 80% or 90% and the significance level is 0.05.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Two-Sample T-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**.

Option	**Value**
Design Tab	
Solve For	**Sample Size**
Alternative Hypothesis	**Two-Sided**
Power	**0.80 0.90**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
d	**0.2 0.5 0.8**

Output

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

Numeric Results

Numeric Results for Two-Sample T-Test
Alternative Hypothesis: H1: d ≠ 0

Target Power	Actual Power	N1	N2	N	Effect Size d	Alpha
0.80	0.8004	393	393	786	0.20	0.050
0.90	0.9002	526	526	1052	0.20	0.050
0.80	0.8015	64	64	128	0.50	0.050
0.90	0.9032	86	86	172	0.50	0.050
0.80	0.8075	26	26	52	0.80	0.050
0.90	0.9015	34	34	68	0.80	0.050

References
Cohen, Jacob. 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates. Hillsdale, New Jersey
Julious, S. A. 2010. Sample Sizes for Clinical Trials. Chapman & Hall/CRC. Boca Raton, FL.
Machin, D., Campbell, M., Tan, B. T., Tan, S. H. 2009. Sample Size Tables for Clinical Studies, 3rd Edition. Wiley-Blackwell.
Ryan, Thomas P. 2013. Sample Size Determination and Power. John Wiley & Sons. New Jersey.

Report Definitions
Target Power is the desired power. May not be achieved because of integer N1 and N2.
Actual Power is the achieved power. Because N1 and N2 are integers, this value is often (slightly) larger than the target power.
N1 and N2 are the number of items sampled from each population.
N is the total sample size, N1 + N2.
Effect Size: $d = (\mu_1 - \mu_2) / \sigma$ is the effect size. Cohen recommended Low = 0.2, Medium = 0.5, and High = 0.8.
Alpha is the probability of rejecting a true null hypothesis.

Summary Statements

Group sample sizes of 393 and 393 achieve 80.04% power to reject the null hypothesis of zero effect size when the population effect size is 0.20 and the significance level (alpha) is 0.050 using a two-sided two-sample equal-variance t-test.

These reports show the values of each of the parameters, one scenario per row.

Chart Section
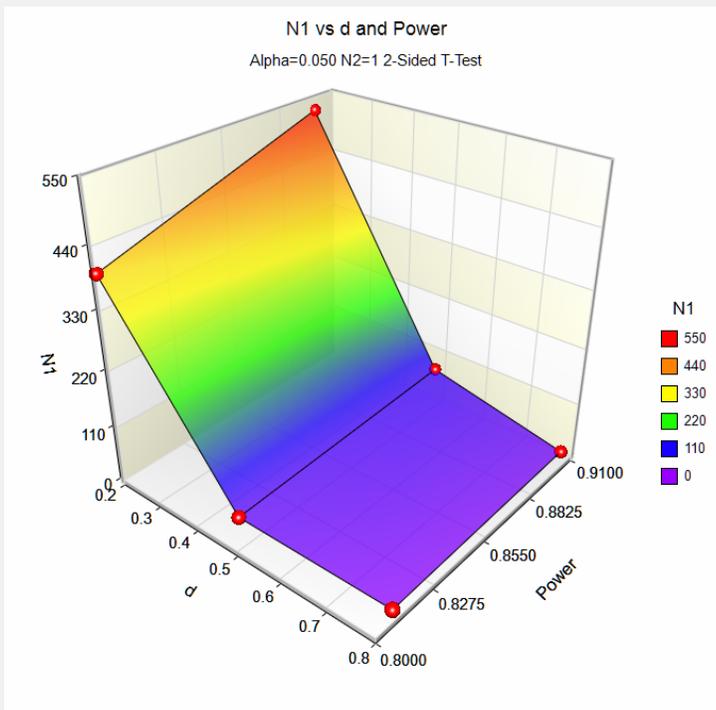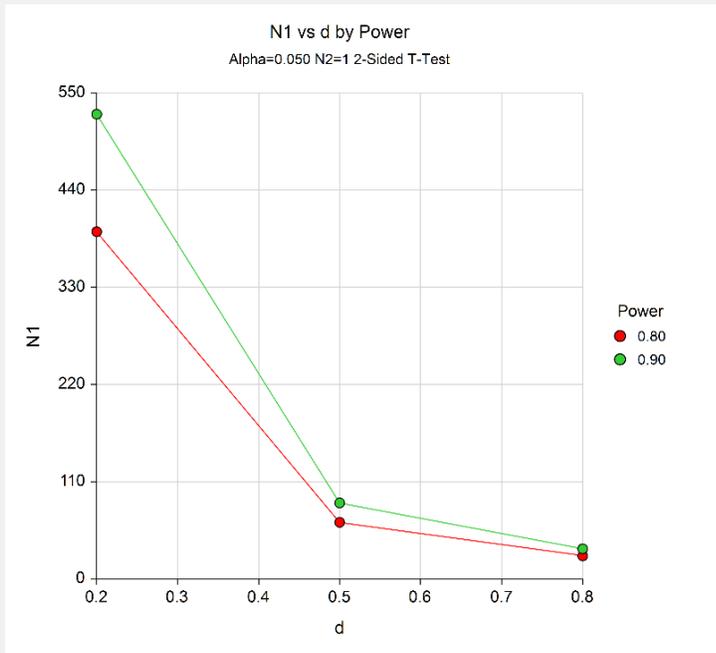
These plots show the relationship between effect size, power, and sample size.

Example 2 – Finding the Power

Researchers wish to compare two types of local anesthesia using a balanced, parallel-group design. Subjects in pain will be randomized to one of two treatment groups, the treatment will be administered, and the subject's evaluation of pain intensity will be measured on a seven-point scale.

The researchers would like to determine the power to detect a small, medium, and large effect size with a two-sided t-test for group sample sizes of 25, 50, 100, 200, 400 and a significance level of 0.05.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Two-Sample T-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**.

Option	Value
Design Tab	
Solve For	**Power**
Alternative Hypothesis	**Two-Sided**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group	**25 50 100 200 400**
d	**0.2 0.5 0.8**

Output

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

Numeric Results

Numeric Results for Two-Sample T-Test
Alternative Hypothesis: H1: $d \neq 0$

Power	N1	N2	N	Effect Size d	Alpha
0.1066	25	25	50	0.20	0.050
0.1677	50	50	100	0.20	0.050
0.2906	100	100	200	0.20	0.050
0.5141	200	200	400	0.20	0.050
0.8073	400	400	800	0.20	0.050
0.4101	25	25	50	0.50	0.050
0.6969	50	50	100	0.50	0.050
0.9404	100	100	200	0.50	0.050
0.9988	200	200	400	0.50	0.050
1.0000	400	400	800	0.50	0.050
0.7915	25	25	50	0.80	0.050
0.9773	50	50	100	0.80	0.050
0.9999	100	100	200	0.80	0.050
1.0000	200	200	400	0.80	0.050
1.0000	400	400	800	0.80	0.050

This report presents the results for various values of *d* and *sample size*.

Example 3 – Validation using Another Procedure

This procedure should give identical results to the **Two-Sample Tests Assuming Equal Variance** procedure when the value of σ there is set to one. We will use this fact to provide a validation problem for this procedure.

If we run that procedure with power = 0.90, alpha = 0.05, $\mu_1 = 1$, $\mu_2 = 0$, $\sigma = 1$, and solve for sample size with N1 = N2, we obtain N1 = N2 = 23.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Two-Sample T-Tests using Effect Size** procedure. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

Option   **Value**

Design Tab

Solve For .... **Sample Size**
Alternative Hypothesis .... **Two-Sided**
Power.. **0.90**
Alpha.. **0.05**
Group Allocation .... **Equal (N1 = N2)**
d **1**

Output

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

Numeric Results

Numeric Results for Two-Sample T-Test
Alternative Hypothesis: H1: d ≠ 0

Target Power	Actual Power	N1	N2	N	Effect Size d	Alpha
0.90	0.9125	23	23	46	1.00	0.050

This procedure also calculated N1 = N2 = 23, thus the procedure is validated.