

Chapter 417

Tests for Paired Means using Effect Size

Introduction

This procedure provides sample size and power calculations for a one- or two-sided paired t-test when the effect size is specified rather than the means and variance. The details of procedure are given in Cohen (1988). In this design, a single population of paired, normally distributed data is sampled and the mean difference is compared to zero by forming the difference scaled by the standard deviation of the differences.

Test Assumptions

When running a paired t-test, the basic assumptions are that the distribution of the paired differences is approximately normal and the subjects are independent.

Test Procedure

If we assume that μ and μ_0 represent the population mean and the specified test value respectively and the (unknown) standard deviation is σ , then 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:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_d^2}{N}}}$$

Tests for Paired Means using Effect Size

This t -statistic follows a t distribution with $N - 1$ 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 *Tests for Paired Means* procedure with the modification that the σ used in that procedure is implicitly set equal to one.

The Effect Size

Suppose we assume that $\mu_1 - \mu_2$ represents the mean of the pairs of the population of interest. If the standard deviation of the pairs 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.

Tests for Paired Means using Effect Size

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 $1 - \text{Beta}$, 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

N (Sample Size)

Enter a value for the sample size (N), the number of individuals in the study. You may enter a single value such as 42, range of values such as 10 to 100 by 10, or a list of values such as 10 30 80 90.

Tests for Paired Means using Effect Size

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 mean and a specified value. The effect size is calculated using

$$d = (\mu_1 - \mu_2) / \sigma$$

where μ_1 is the mean assumed by the alternative hypothesis for the first paired variable, μ_2 is the mean assumed by the alternative hypothesis for the second paired variable, and σ is your estimate of the population standard deviation of the differences.

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.

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

Tests for Paired Means using Effect Size

Example 1 – Finding the Sample Size

Researchers wish to compare a two treatments for chronic pain. Subjects suffering from chronic pain are given on treatment on one night and the other treatment on the following night. The treatment to be given first is selected by a coin-toss. 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, paired 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 **Tests for Paired Means 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>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided
Power	0.8 0.9
Alpha	0.05
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 One-Sample T-Test**

Alternative Hypothesis: $H_1: d \neq 0$

Target Power	Actual Power	N	Effect Size d	Alpha
0.80	0.8017	199	0.20	0.050
0.90	0.9004	265	0.20	0.050
0.80	0.8078	34	0.50	0.050
0.90	0.9000	44	0.50	0.050
0.80	0.8213	15	0.80	0.050
0.90	0.9092	19	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 N.

Actual Power is the achieved power. Because N is an integer, this value is often (slightly) larger than the target power.

N is the number of items sampled from the population.

Effect Size: $d = (\mu_1 - \mu_2) / \sigma$ is the effect size. Cohen recommended Low = 0.2, Medium = 0.5, and High = 0.8.

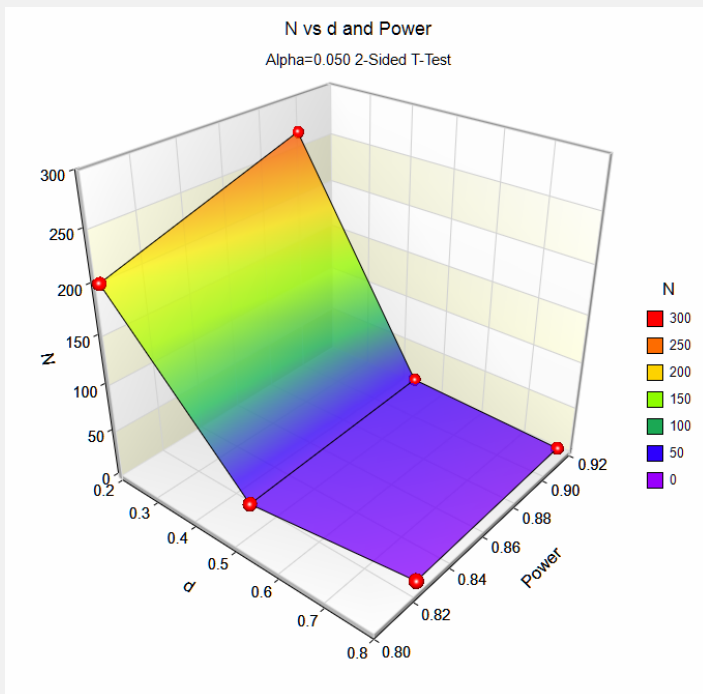
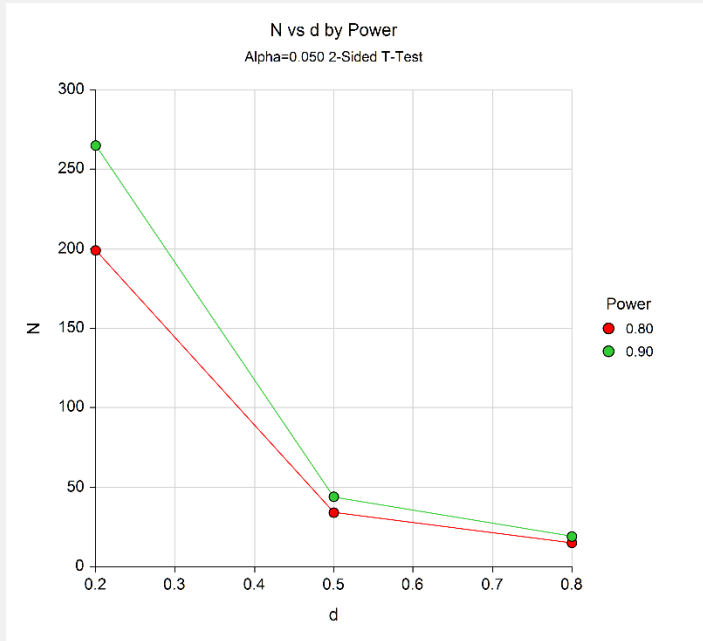
Tests for Paired Means using Effect Size

Summary Statements

A sample size of 199 data pairs achieves 80.2% 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 paired t-test.

This report shows the values of each of the parameters, one scenario per row.

Plots Section



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

Tests for Paired Means using Effect Size

Example 2 – Validation using Another Procedure

This procedure should give identical results to the **Tests for Paired Means** 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, Mean of Paired Differences = 0.5, SD = 1, and solve for sample size. The result is N = 44.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Paired Means 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>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided
Power	0.9
Alpha	0.05
d	0.5

Output

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

Numeric Results**Numeric Results for Paired T-Test**Alternative Hypothesis: H1: $d \neq 0$

Target Power	Actual Power	N	Effect Size d	Alpha
0.90	0.9000	44	0.50	0.050

This procedure also calculated N = 44, thus the procedure is validated.