Chapter 518

One-Sample T-Tests for Equivalence

Introduction

This procedure allows you to study the power and sample size of *t*-tests of equivalence of one mean with a value (e.g. a historical control or gold standard). Schuirmann's (1987) two one-sided tests (TOST) approach is used to test equivalence. The one-sample *t*-test is commonly used in this situation.

Outline of an Equivalence Test

PASS follows the *two one-sided tests* approach described by Schuirmann (1987) and Phillips (1990). It will be convenient to adopt the following specialize notation for the discussion of these tests.

Parameter	PASS Input/Output	Interpretation
E_L, E_U	EL, EU	*Lower* and *Upper Equivalence Limits*. If the actual mean is between E_L and E_U, the treatment is said to be *equivalent* to the reference value.
μ	μ	*Population mean*. If the data are paired differences, this is the mean of those differences. This parameter will be estimated by the study.
μ_1	$\mu1$	*Actual population mean at which power is calculated*. This is the assumed population mean used in all calculations.

With $E_L < E_U$, the null hypothesis of non-equivalence is

$$H_0: \mu \leq E_L \ \text{ or } \ \mu \geq E_U.$$

The alternative hypothesis of equivalence is

$$H_1: E_L < \mu < E_U.$$

One-Sample T-Test Statistic

A one-sample *t*-test is used to analyze the data. The test assumes that the data are a simple random sample from a population of normally-distributed values that have the same variance. This assumption implies that the observations are continuous and normal. The calculation of the two, one-sided *t*-tests proceeds as follows

$$t_L = \frac{\bar{x} - E_L}{s_x/\sqrt{N}}$$

$$t_U = \frac{\bar{x} - E_U}{s_x/\sqrt{N}}$$

where s_x is the sample standard deviation of the measurements. The test is usually calculated using a $100(1 - 2\alpha)\%$ confidence interval of the mean. If both limits of this confidence interval are between E_L and E_U, equivalence is concluded.

Power Calculation

The power of this test is

$$\Pr\left(t_L \geq t_{1-\alpha,N-1} \text{ and } t_U \leq t_{\alpha,N-1}\right)$$

where t_L and t_U are distributed as the bivariate, noncentral t distribution with noncentrality parameters Δ_L and Δ_U given by

$$\Delta_L = \frac{\mu - E_L}{\sigma\sqrt{1/N}}$$

$$\Delta_U = \frac{\mu - E_U}{\sigma\sqrt{1/N}}.$$

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 the parameters associated with this test such as the means, sample sizes, alpha, and power.

Solve For

Solve For

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

Select *Sample Size* when you want to calculate the sample size needed to achieve a given power and alpha level. Select *Power* when you want to calculate the power of an experiment that has already been run.

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. 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. 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

N (Sample Size)

This option specifies one or more values of the sample size, the number of individuals in the study (total subjects). This value must be an integer greater than one.

You may enter a list of values using the syntax *50,100,150,200,250* or *50 to 250 by 50*.

Effect Size – Equivalence Limits

EU (Upper Equivalence Limit)

Enter the upper limit on equivalence. Mean differences outside EL and EU are not considered equivalent. Differences between them are considered equivalent.

Note that you must have $EL < \mu 1 < EU$.

EL (Lower Equivalence Limit)

Enter the lower limit on equivalence. Differences outside EL and EU are not considered equivalent. Differences between them are.

If you want symmetric limits about zero, enter "-Upper Limit" for EL to force EL = -|EU|.

Note that you must have $EL < \mu 1 < EU$. Finally, the scale of these numbers must match the scale of σ.

Effect Size – Mean

μ1 (Actual Mean)

Enter one or more values for the actual value of the mean at which power and sample size will be calculated.

This value can be any number. For paired data, this value is often set to zero, but it can be non-zero as long as $EL < \mu 1 < EU$.

Effect Size – Standard Deviation

σ (Standard Deviation)

Enter one or more values for the standard deviation σ of the data values. You can press the σ button to load the Standard Deviation Estimator window.

Example 1 – Finding Power

A one-group design is to be used to compare the impact on diastolic blood pressure of a new drug with the known standard drug. The average diastolic blood pressure after administration of the reference drug is 96 mmHg. Researchers believe this average may drop to 92 mmHg with the use of a new drug. An estimate of σ is 25.

Following FDA guidelines, the researchers want to show that the diastolic blood pressure with the new drug is within 20% of the diastolic blood pressure with the standard drug. Thus, the equivalence limits of the mean difference of the two drugs are 96 - 19.2 = 76.8 and 96 + 19.2 = 115.2. They decide to calculate the power for a range of sample sizes between 5 and 50. 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 **One-Sample T-Tests for Equivalence** procedure window by expanding **Means**, then **One Mean**, selecting **Equivalence**, and then clicking on **One-Sample T-Tests for Equivalence**. 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	**Power**
Alpha...	**0.05**
N (Sample Size).....................................	**5 10 15 20 30 40 50**
EU (Upper Equivalence Limit)	**115.2**
EL (Lower Equivalence Limit).................	**76.8**
μ1 (Actual Mean)	**92**
σ (Standard Deviation)	**25**

Annotated Output

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

Numeric Results

Numeric Results ——
Hypotheses: H0: μ ≤ EL or μ ≥ EU   vs.   H1: EL < μ < EU

Power	Sample Size N	Lower Equiv Limit EL	Upper Equiv Limit EU	Actual Mean μ1	Standard Deviation σ	Alpha
0.11327	5	76.800	115.200	92.000	25.000	0.050
0.41782	10	76.800	115.200	92.000	25.000	0.050
0.68518	15	76.800	115.200	92.000	25.000	0.050
0.82597	20	76.800	115.200	92.000	25.000	0.050
0.94542	30	76.800	115.200	92.000	25.000	0.050
0.98350	40	76.800	115.200	92.000	25.000	0.050
0.99526	50	76.800	115.200	92.000	25.000	0.050

References
Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, Third
   Edition. Taylor & Francis/CRC. Boca Raton, Florida.
Mathews, Paul. 2010. Sample Size Calculations - Practical Methods for Engineers and Scientists. Mathews Malnar
   and Bailey. Fairport Harbor, OH.
Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York.
   Volume 2, 1367-1372.

Report Definitions
Power is the probability of rejecting the null hypothesis when it is false. It should be close to one.
N is the sample size, the number of subjects (or pairs) in the study.
EL is the minimum allowable mean that still results in equivalence.
EU is the maximum allowable mean that still result in equivalence.
$\mu 1$ is the actual value of the mean at which the power and sample size are computed.
σ is the standard deviation of the response (or standard deviation of differences for paired data). It
   measures the variability in the population.
Alpha is the probability of rejecting the null hypothesis when it is true. It should be small.

Summary Statements ───
In an equivalence test of one mean using two one-sided t-tests, a sample size of 5 achieves 11%
power at a 5.0% significance level when the actual mean is 92.000, the estimated standard
deviation is 25.000, and the equivalence limits are 76.800 and 115.200.

This report shows the power for the indicated scenarios. Note that if they want 90% power, they will require a
sample of around 30 subjects.

Plots Section

Power vs N

$\mu 1 = 92.00$ $\sigma = 25.00$ EL=76.80 EU=115.20 Alpha=0.050 Equivalence T-Test
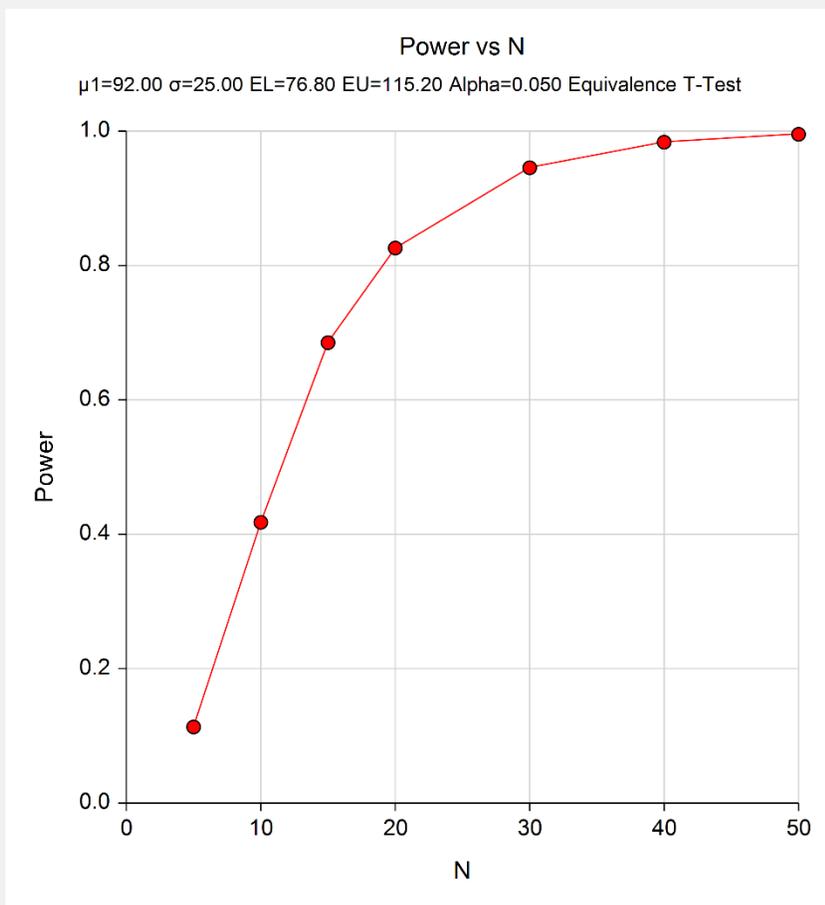
This plot shows the power versus the sample size.

Example 2 – Validation using Chow, Shao, Wang, and Lokhnygina (2018)

Chow, Shao, Wang, and Lokhnygina (2018) presents an example on pages 46 and 47 of a one-sample equivalence *t*-test sample size calculation in which EU = 0.05, EL = -0.05, µ1 = 0.0, σ = 0.1, alpha = 0.05, and power = 0.80. They obtain a sample size of 36.

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-Sample T-Tests for Equivalence** procedure window by expanding **Means**, then **One Mean**, selecting **Equivalence**, and then clicking on **One-Sample T-Tests for Equivalence**. 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	**Sample Size**
Power..	**0.80**
Alpha..	**0.05**
EU (Upper Equivalence Limit)	**0.05**
EL (Lower Equivalence Limit).................	**-0.05**
µ1 (Actual Mean)	**0**
σ (Standard Deviation)	**0.1**

Output

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

Numeric Results ───
Hypotheses: H0: µ ≤ EL or µ ≥ EU   vs.   H1: EL < µ < EU

Power	Sample Size N	Lower Equiv Limit EL	Upper Equiv Limit EU	Actual Mean µ1	Standard Deviation σ	Alpha
0.80515	36	-0.050	0.050	0.000	0.100	0.050

The sample size of 36 matches Chow, Shao, Wang, and Lokhnygina (2018) exactly.

Example 3 – Validation using Phillips (1990)

Phillips (1990) page 142 presents a table of sample sizes for various parameter values. In this table, the treatment mean, standard deviation, and equivalence limits are all specified as percentages of the reference mean. We will reproduce the second line of the table in which the square root of the within mean square error is 20% (σ of 28.284%); the equivalence limits are 80% and 120%; the treatment mean is 100%, 95%, 90%, and 85%; the power is 70%; and the significance level is 0.05.

Phillips reports total sample sizes of 16, 20, 40, and 152 corresponding to the four treatment mean percentages. We will now setup this example in **PASS**.

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-Sample T-Tests for Equivalence** procedure window by expanding **Means**, then **One Mean**, selecting **Equivalence**, and then clicking on **One-Sample T-Tests for Equivalence**. 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**
Power... **0.70**
Alpha.. **0.05**
EU (Upper Equivalence Limit) **120**
EL (Lower Equivalence Limit)................. **80**
$\mu 1$ (Actual Mean) **85 90 95 100**
σ (Standard Deviation) **28.284**

Output

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

Numeric Results

Numeric Results ————————————————————————————————————
Hypotheses: H0: $\mu \le$ EL or $\mu \ge$ EU   vs.   H1: EL $< \mu <$ EU

Power	Sample Size N	Lower Equiv Limit EL	Upper Equiv Limit EU	Actual Mean $\mu 1$	Standard Deviation σ	Alpha
0.70015	152	80.000	120.000	85.000	28.284	0.050
0.70958	40	80.000	120.000	90.000	28.284	0.050
0.72396	20	80.000	120.000	95.000	28.284	0.050
0.70750	16	80.000	120.000	100.000	28.284	0.050

PASS obtains the same samples sizes as Phillips (1990).