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	$\mu 1$	*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 – 2α)% 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

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

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

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

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**
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**

Output

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

Numeric Reports

Numeric Results

Solve For: Power
Hypotheses: H0: μ ≤ EL or μ ≥ EU vs. H1: EL < μ < EU

		Equivalence Limits				
Power	Sample Size N	Lower EL	Upper EU	Actual Mean μ1	Standard Deviation σ	Alpha
0.11327	5	76.8	115.2	92	25	0.05
0.41782	10	76.8	115.2	92	25	0.05
0.68518	15	76.8	115.2	92	25	0.05
0.82597	20	76.8	115.2	92	25	0.05
0.94542	30	76.8	115.2	92	25	0.05
0.98350	40	76.8	115.2	92	25	0.05
0.99526	50	76.8	115.2	92	25	0.05

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N The sample size, the number of subjects (or pairs) in the study.
EL The minimum allowable mean that still results in equivalence.
EU The maximum allowable mean that still result in equivalence.
μ1 The actual value of the mean at which the power and sample size are computed.
σ The standard deviation of the response (or standard deviation of differences for paired data). It measures the variability in the population.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A single-group design will be used to test equivalence of a mean to a value, by testing the mean against the lower and upper equivalence limits of 76.8 and 115.2 (H0: μ ≤ 76.8 or μ ≥ 115.2 versus H1: 76.8 < μ < 115.2). The comparison will be made using two one-sided t-tests, with an overall Type I error rate (α) of 0.05. The underlying standard deviation is assumed to be 25. To detect a mean of 92 with a sample size of 5, the power is 0.11327.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	5	7	2
20%	10	13	3
20%	15	19	4
20%	20	25	5
20%	30	38	8
20%	40	50	10
20%	50	63	13

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, 7 subjects should be enrolled to obtain a final sample size of 5 subjects.

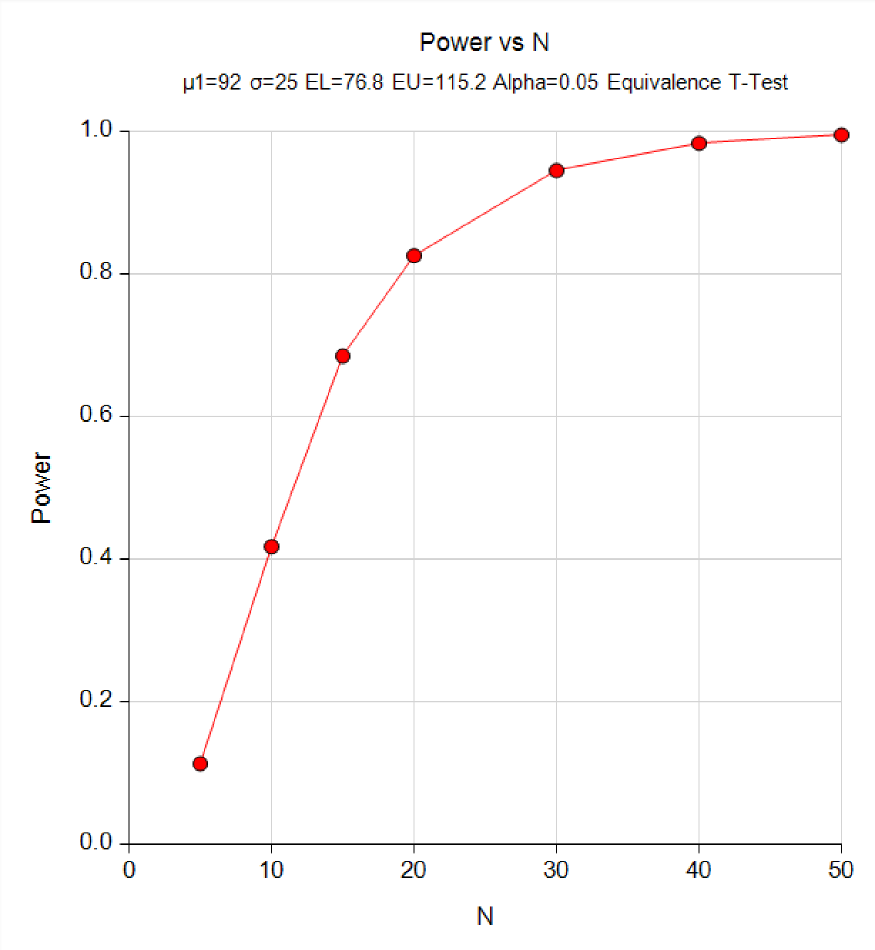
References

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.

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

Plots



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

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... **0.80**
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

Solve For: Sample Size
Hypotheses: H0: μ ≤ EL or μ ≥ EU vs. H1: EL < μ < EU

Power	Sample Size N	Equivalence Limits Lower EL	Equivalence Limits Upper EU	Actual Mean μ1	Standard Deviation σ	Alpha
0.80515	36	-0.05	0.05	0	0.1	0.05

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

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 .... **Sample Size**
Power.. **0.70**
Alpha.. **0.05**
EU (Upper Equivalence Limit)........................ **120**
EL (Lower Equivalence Limit) **80**
μ1 (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

Solve For: Sample Size
Hypotheses: H0: μ ≤ EL or μ ≥ EU vs. H1: EL < μ < EU

Power	Sample Size N	Equivalence Limits Lower EL	Equivalence Limits Upper EU	Actual Mean μ1	Standard Deviation σ	Alpha
0.70015	152	80	120	85	28.284	0.05
0.70958	40	80	120	90	28.284	0.05
0.72396	20	80	120	95	28.284	0.05
0.70750	16	80	120	100	28.284	0.05

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