

## Chapter 413

# Superiority by a Margin Tests for One Mean

---

### Introduction

This module computes power and sample size for tests in one-sample designs with a superiority margin in which the outcome is distributed as a normal random variable. This includes the analysis of the differences between paired values.

The details of sample size calculation for the one-sample design are presented in the Inequality Tests for One Mean chapter and they will not be duplicated here. This chapter only discusses those changes necessary for superiority (or non-zero null) tests. Sample size formulas for non-inferiority and superiority hypothesis tests of a single mean are presented in Chow et al. (2003) page 50.

The *one-sample t-test* is used to test whether a population mean is different from a specific value. When the data are differences between paired values, this test is known as the *paired t-test*. This module also calculates the power of the non-parametric analog of the *t-test*, the *Wilcoxon test*.

---

### Paired Designs

Paired data may occur because two measurements are made on the same subject or because measurements are made on two subjects that have been matched according to other variables. Hypothesis tests on paired data can be analyzed by considering the difference between the paired items as the response. The distribution of differences is usually symmetric. In fact, the distribution must be symmetric if the individual distributions of the two items are identical. Hence, the paired *t-test* and the Wilcoxon signed-rank test are appropriate for paired data even when the distributions of the individual items are not normal.

In paired designs, the variable of interest is the difference between two individual measurements. Although the non-inferiority hypothesis refers to the difference between two individual means, the actual values of those means are not needed. All that is needed is their difference.

---

### The Statistical Hypotheses

Both non-inferiority and superiority tests are examples of directional (one-sided) tests and their power and sample size could be calculated using the One-Sample T-Test procedure. However, at the urging of our users, we have developed this module which provides the input and output options that are convenient for superiority hypothesis tests. This section will review the specifics of superiority testing.

## Superiority by a Margin Tests for One Mean

Remember that in the usual t-test setting, the null ( $H_0$ ) and alternative ( $H_1$ ) hypotheses for one-sided tests are defined as

$$H_0: \mu_x \leq A \text{ versus } H_1: \mu_x > A$$

Rejecting  $H_0$  implies that the mean is larger than the value  $A$ . This test is called an *upper-tail test* because  $H_0$  is rejected in samples in which the sample mean is larger than  $A$ .

Following is an example of a *lower-tail test*.

$$H_0: \mu_x \geq A \text{ versus } H_1: \mu_x < A$$

*Superiority* tests are special cases of the above directional tests. It will be convenient to adopt the following specialize notation for the discussion of these tests.

<u>Parameter</u>	<u>PASS Input/Output</u>	<u>Interpretation</u>
$\mu_T$	Not used	<i>Population mean.</i> If the data are paired differences, this is the mean of those differences.
$\mu_R$	Not used	<i>Reference value.</i> Usually, this is the mean of a reference population. If the data are paired differences, this is the hypothesized value of the mean difference.
$M_s$	SM	<i>Margin of superiority.</i> This is a tolerance value that defines the magnitude of difference that is required for practical importance. This may be thought of as the smallest difference from the reference value that is considered to be different.
$\delta$	D	<i>True difference.</i> This is the value of $\mu_T - \mu_R$ , the difference between the mean and the reference value, at which the power is calculated.

Note that the actual values of  $\mu_T$  and  $\mu_R$  are not needed. Only their difference is needed for power and sample size calculations.

---

## Superiority Tests

A *superiority by a margin test* tests that the mean is better than that of the baseline (reference) population by more than a small superiority margin. The actual direction of the hypothesis depends on the whether higher values of the response are good or bad.

### Case 1: High Values Good

In this case, higher values are better. The hypotheses are arranged so that rejecting the null hypothesis implies that the mean is greater than the reference value by at least the margin of superiority. The value of  $\delta$  must be greater than  $|M_s|$ . Equivalent sets of the null and alternative hypotheses are

$$H_0: \mu_T \leq \mu_R + |M_s| \text{ versus } H_1: \mu_T > \mu_R + |M_s|$$

$$H_0: \mu_T - \mu_R \leq |M_s| \text{ versus } H_1: \mu_T - \mu_R > |M_s|$$

$$H_0: \delta \leq |M_s| \text{ versus } H_1: \delta > |M_s|$$

## Superiority by a Margin Tests for One Mean

### Case 2: High Values Bad

In this case, lower values are better. The hypotheses are arranged so that rejecting the null hypothesis implies that the mean is less than the reference value by at least the margin of superiority. The value of  $\delta$  must be less than  $-|M_S|$ . Equivalent sets of the null and alternative hypotheses are

$$H_0 : \mu_T \geq \mu_R - |M_S| \quad \text{versus} \quad H_1 : \mu_T < \mu_R - |M_S|$$

$$H_0 : \mu_T - \mu_R \geq -|M_S| \quad \text{versus} \quad H_1 : \mu_T - \mu_R < -|M_S|$$

$$H_0 : \delta \geq -|M_S| \quad \text{versus} \quad H_1 : \delta < -|M_S|$$

### Example

A superiority test example will set the stage for the discussion of the terminology that follows. Suppose that a test is to be conducted to determine if a new cancer treatment substantially improves the mean bone density. The adjusted mean bone density (AMBD) in the population of interest is 0.002300 gm/cm with a standard deviation of 0.000300 gm/cm. Clinicians decide that if the treatment increases AMBD by more than 5% (0.000115 gm/cm), it provides a significant health benefit.

The hypothesis of interest is whether the AMBD in the treated group is greater than  $0.002300 + 0.000115 = 0.002415$ . The statistical test will be set up so that if the null hypothesis that the AMBD is greater than or equal to 0.002415 is rejected, the conclusion will be that the new treatment is superior, at least in terms of AMBD. The value 0.000115 gm/cm is called the *margin of superiority*.

### Test Statistics

This section describes the test statistics that are available in this procedure.

#### One-Sample T-Test

The one-sample t-test assumes that the data are a simple random sample from a population of normally-distributed values that all have the same mean and variance. This assumption implies that the data are continuous and their distribution is symmetric. The calculation of the t-test proceeds as follow

$$t_{n-1} = \frac{\bar{X} - D0}{s_{\bar{X}}}$$

where

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n},$$

$$s_{\bar{X}} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}},$$

and  $D0$  is the value of the mean hypothesized by the null hypothesis.

The significance of the test statistic is determined by computing the p-value. If this p-value is less than a specified level (usually 0.05), the hypothesis is rejected. Otherwise, no conclusion can be reached.

## Superiority by a Margin Tests for One Mean

### Wilcoxon Signed-Rank Test

The Wilcoxon signed-rank test is a popular, nonparametric substitute for the t-test. It assumes that the data follow a symmetric distribution. The test is computed using the following steps.

1. Subtract the hypothesized mean,  $D0$ , from each data value. Rank the values according to their absolute values.
2. Compute the sum of the positive ranks  $Sp$  and the sum of the negative ranks  $Sn$ . The test statistic,  $W$ , is the minimum of  $Sp$  and  $Sn$ .
3. Compute the mean and standard deviation of  $W$  using the formulas

$$\mu_{W_n} = \frac{n(n+1)}{4} \quad \text{and} \quad \sigma_{W_n} = \sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum t^3 - \sum t}{48}}$$

where  $t$  represents the number of times the  $i$ th value occurs.

4. Compute the  $z$  value using

$$z_W = \frac{W - \mu_{W_n}}{\sigma_{W_n}}$$

The significance of the test statistic is determined by computing the p-value using the standard normal distribution. If this p-value is less than a specified level (usually 0.05), the null hypothesis is rejected in favor of the alternative hypothesis. Otherwise, no conclusion can be reached.

## Computing the Power

The power is calculated as follows for a directional alternative (one-tailed test) in which  $D1 > D0$ .  $D1$  is the value of the mean at which the power is computed.

1. Find  $t_\alpha$  such that  $1 - T_{n-1}(t_\alpha) = \alpha$ , where  $T_{n-1}(t_\alpha)$  is the area to the left of  $x$  under a central-t curve with  $n - 1$  degrees of freedom.
2. Calculate  $x_a = D0 + t_\alpha \frac{\sigma}{\sqrt{n}}$ .
3. Calculate the noncentrality parameter  $\lambda = \frac{D1 - D0}{\frac{\sigma}{\sqrt{n}}}$ .
4. Calculate  $t_a = \frac{x_a - D1}{\frac{\sigma}{\sqrt{n}}} + \lambda$ .
5. Calculate the power =  $1 - T'_{n-1,\lambda}(t_a)$ , where  $T'_{n-1,\lambda}(x)$  is the area to the left of  $x$  under a noncentral-t curve with degrees of freedom  $n - 1$  and noncentrality parameter  $\lambda$ .

---

## 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 will be of interest.

---

#### Solve For

##### Solve For

This option specifies the parameter to be calculated from the values of the other parameters. Under most conditions, you would select either *Power* or *Sample Size*.

Select *Sample Size* when you want to determine the sample size needed to achieve a given power and alpha error level.

Select *Power* when you want to calculate the power.

---

#### Test

##### Higher Means Are

This option defines whether higher values of the response variable are to be considered better or worse. The choice here determines the direction of the test.

If Higher Means Are Better the null hypothesis is  $\text{Diff} \leq \text{SM}$  and the alternative hypothesis is  $\text{Diff} > \text{SM}$ . If Higher Means Are Worse the null hypothesis is  $\text{Diff} \geq -\text{SM}$  and the alternative hypothesis is  $\text{Diff} < -\text{SM}$ .

##### Nonparametric Adjustment

This option makes appropriate sample size adjustments for the Wilcoxon test. Results by Al-Sunduqchi and Guenther (1990) indicate that power calculations for the Wilcoxon test may be made using the standard  $t$  test formulations with a simple adjustment to the sample size. The size of the adjustment depends upon the actual distribution of the data. They give sample size adjustment factors for four distributions. These are 1 for the uniform distribution,  $2/3$  for the double exponential distribution,  $9 / \pi^2$  for the logistic distribution, and  $\pi / 3$  for the normal distribution.

The options are as follows:

- **Ignore**  
Do not make a Wilcoxon adjustment. This indicates that you want to analyze a  $t$  test, not the Wilcoxon test.
- **Uniform**  
Make the Wilcoxon sample size adjustment assuming the uniform distribution. Since the factor is one, this option performs the same as Ignore. It is included for completeness.
- **Double Exponential**  
Make the Wilcoxon sample size adjustment assuming that the data actually follow the double exponential distribution.

## Superiority by a Margin Tests for One Mean

- **Logistic**  
Make the Wilcoxon sample size adjustment assuming that the data actually follow the logistic distribution.
- **Normal**  
Make the Wilcoxon sample size adjustment assuming that the data actually follow the normal distribution.

### Population Size

This is the number of subjects in the population. Usually, you assume that samples are drawn from a very large (infinite) population. Occasionally, however, situations arise in which the population of interest is of limited size. In these cases, appropriate adjustments must be made.

When a finite population size is specified, the standard deviation is reduced according to the formula

$$\sigma_1^2 = \left(1 - \frac{n}{N}\right)\sigma^2$$

where  $n$  is the sample size,  $N$  is the population size,  $\sigma$  is the original standard deviation, and  $\sigma_1$  is the new standard deviation.

The quantity  $n/N$  is often called the sampling fraction. The quantity  $\left(1 - \frac{n}{N}\right)$  is called the *finite population correction factor*.

---

## 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*.

## Superiority by a Margin Tests for One Mean

---

### Sample Size

#### **N (Sample Size)**

This option specifies one or more values of the sample size, the number of individuals in the study. 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 – Mean Difference

#### **SM (Superiority Margin)**

This is the magnitude of the margin of superiority. It must be entered as a positive number.

When higher means are better, this value is the distance above the reference value that is required to be considered superior. When higher means are worse, this value is the distance below the reference value that is required to be considered superior.

#### **D (True Difference)**

This is the actual difference between the mean and the reference value.

When higher means are better, this value should be greater than SM. When higher means are worse, this value should be negative and greater in magnitude than SM.

---

### Effect Size – Standard Deviation

#### **Standard Deviation**

This option specifies one or more values of the standard deviation. This must be a positive value. *PASS* includes a special module for estimating the standard deviation. This module may be loaded by pressing the *SD* button. Refer to the Standard Deviation Estimator chapter for further details.

## Superiority by a Margin Tests for One Mean

### Example 1 – Power Analysis

Suppose that a test is to be conducted to determine if a new cancer treatment improves the mean bone density. The adjusted mean bone density (AMBD) in the population of interest is 0.002300 gm/cm with a standard deviation of 0.000300 gm/cm. Clinicians decide that if the treatment increases AMBD by more than 5% (0.000115 gm/cm), it generates a significant health benefit. They also want to consider what would happen if the margin of superiority is set to 2.5% (0.0000575 gm/cm).

The analysis will be a superiority test using the t-test at the 0.025 significance level. Power is to be calculated assuming that the new treatment has 7.5% improvement on AMBD. Several sample sizes between 20 and 300 will be analyzed. The researchers want to achieve a power of at least 90%. All numbers have been multiplied by 10000 to make the reports and plots easier to read.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Superiority by a Margin Tests for One Mean** procedure window by expanding **Means**, then **One Mean**, then clicking on **Superiority by a Margin**, and then clicking on **Superiority by a Margin Tests for One Mean**. 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
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Higher Means Are.....	<b>Better</b>
Nonparametric Adjustment .....	<b>Ignore</b>
Population Size.....	<b>Infinite</b>
Alpha.....	<b>0.025</b>
N (Sample Size).....	<b>20 40 60 80 100 150 200 300</b>
SM (Superiority Margin).....	<b>0.575 1.15</b>
D (True Difference).....	<b>1.725</b>
S (Standard Deviation) .....	<b>3</b>

### Annotated Output

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

Numeric Results for Superiority Test (H0: Diff ≤ SM; H1: Diff > SM)						
Higher Means are Better						
Test Statistic: T-Test						
	Superiority	Actual	Significance		Standard	
Power	N	Margin	Difference	Level	Beta	Deviation
		(SM)	(D)	(Alpha)		(S)
0.36990	20	0.575	1.725	0.02500	0.63010	3.000
0.65705	40	0.575	1.725	0.02500	0.34295	3.000
0.83164	60	0.575	1.725	0.02500	0.16836	3.000
0.92317	80	0.575	1.725	0.02500	0.07683	3.000
0.96682	100	0.575	1.725	0.02500	0.03318	3.000
0.99658	150	0.575	1.725	0.02500	0.00342	3.000
0.99970	200	0.575	1.725	0.02500	0.00030	3.000
1.00000	300	0.575	1.725	0.02500	0.00000	3.000
0.12601	20	1.150	1.725	0.02500	0.87399	3.000
(report continues)						

### Superiority by a Margin Tests for One Mean

**Report Definitions**

H0 (null hypothesis) is  $Diff \leq SM$ , where  $Diff = Mean - Reference Value$  for the one-sample case, and  $Diff = Mean of Differences - Reference Value$  for the paired case.

H1 (alternative hypothesis) is  $Diff > SM$ .

Power is the probability of rejecting H0 when it is false.

N is the sample size, the number of subjects in the study.

SM is the magnitude of the margin of superiority. Since higher means are better, this value is positive and is the distance above the reference value that is required to be considered superior.

D is the mean difference (treatment - reference value) at which the power is computed.

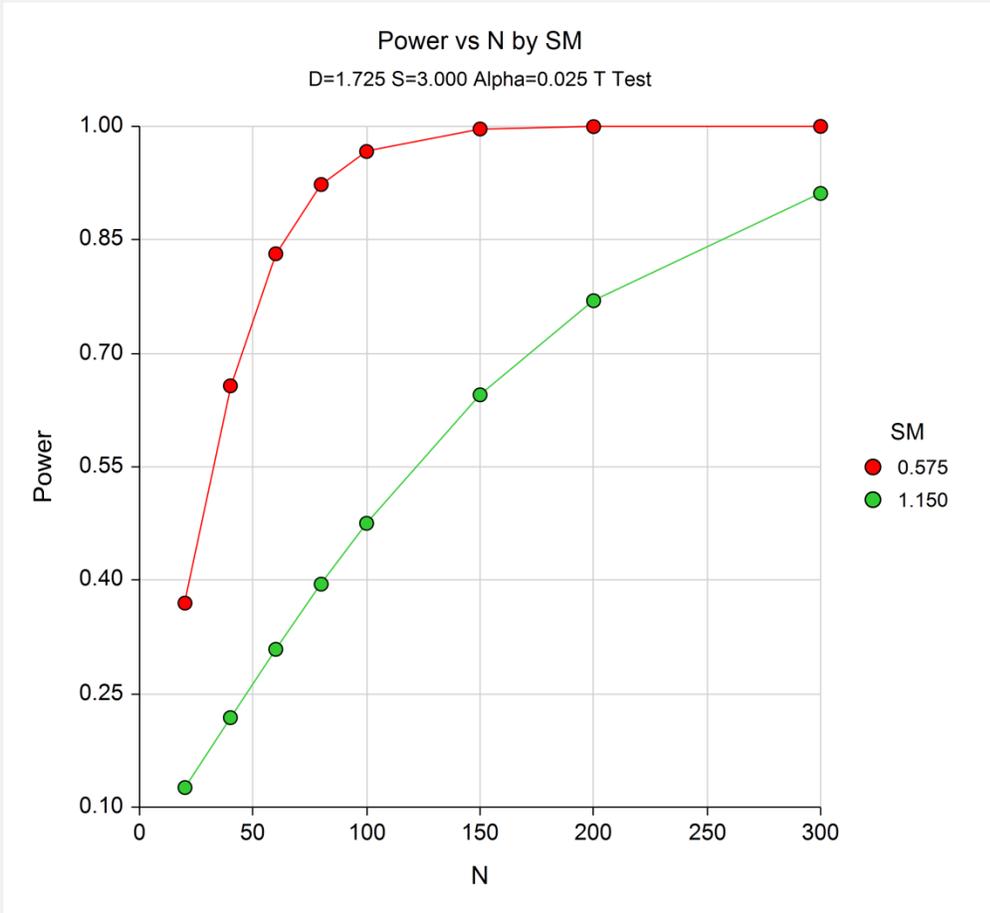
Alpha is the probability of rejecting H0 when it is true, which is the probability of a false positive.

Beta is the probability of accepting H0 when it is false, which is the probability of a false negative.

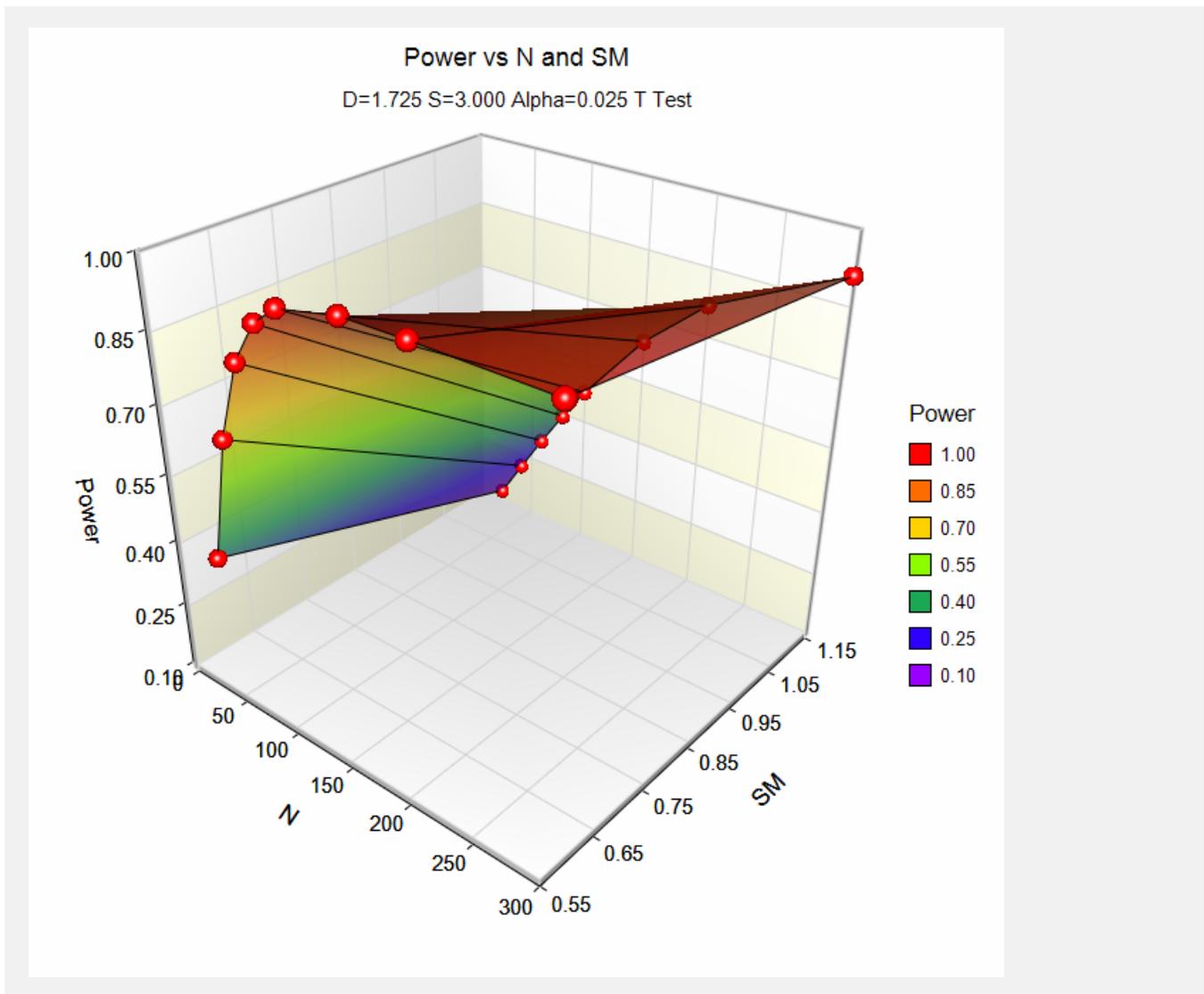
S is the standard deviation of the response. It measures the variability in the population.

**Summary Statements**

A sample size of 20 achieves 37% power to detect superiority using a one-sided t-test when the margin of superiority is 0.575 and the true difference between the mean and the reference value is 1.725. The data are drawn from a single population with a standard deviation of 3.000. The significance level (alpha) of the test is 0.02500.



## Superiority by a Margin Tests for One Mean



The above report shows that for  $SM = 1.15$ , the sample size necessary to obtain 90% power is just under 300. However, if  $SM = 0.575$ , the required sample size is only about 75.

## Superiority by a Margin Tests for One Mean

### Example 2 – Finding the Sample Size

Continuing with Example 1, the researchers want to know the exact sample size for each value of SM to achieve 90% power.

#### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Superiority by a Margin Tests for One Mean** procedure window by expanding **Means**, then **One Mean**, then clicking on **Superiority by a Margin**, and then clicking on **Superiority by a Margin Tests for One Mean**. 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
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Higher Means Are .....	<b>Better</b>
Nonparametric Adjustment .....	<b>Ignore</b>
Population Size .....	<b>Infinite</b>
Power .....	<b>0.90</b>
Alpha .....	<b>0.025</b>
SM (Superiority Margin) .....	<b>0.575 1.15</b>
D (True Difference) .....	<b>1.725</b>
S (Standard Deviation) .....	<b>3</b>

#### Output

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

#### Numeric Results

**Numeric Results for Superiority Test (H0: Diff ≤ SM; H1: Diff > SM)**  
**Higher Means are Better**  
**Test Statistic: T-Test**

		Superiority Margin	Actual Difference	Significance Level		Standard Deviation
Power	N	(SM)	(D)	(Alpha)	Beta	(S)
0.90215	74	0.575	1.725	0.02500	0.09785	3.000
0.90005	288	1.150	1.725	0.02500	0.09995	3.000

This report shows the exact sample size requirement for each value of SM.

### Example 3 – Validation

This procedure uses the same mechanics as the Non-Inferiority Tests for One Mean procedure. We refer the user to Example 3 of Chapter 415 for the validation.