

## Chapter 410

# Tests for One Mean (Simulation)

---

### Introduction

This procedure allows you to study the power and sample size of several statistical tests of the hypothesis that the population mean is equal to a specific value versus the alternative that it is greater than, less than, or not equal to that value. The one-sample t-test is commonly used in this situation, but other tests have been developed for situations where the data are not normally distributed. These additional tests include the Wilcoxon signed-rank test, the sign test, and the computer-intensive bootstrap test. When the population follows the exponential distribution, a test based on this distribution should be used.

The t-test assumes that the data are normally distributed. When this assumption does not hold, the t-test is still used hoping that its robustness will produce accurate results. This procedure allows you to study the accuracy of various tests using simulation techniques. A wide variety of distributions can be simulated to allow you to assess the impact of various forms of non-normality on each test's accuracy.

The details of the power analysis of the t-test using analytic techniques are presented in the **PASS** chapter entitled "One-Sample T-Tests" and will not be duplicated here. This chapter will be confined to power analysis using computer simulation.

---

### Technical Details

*Computer simulation* allows one to estimate the power and significance level that is actually achieved by a test procedure in situations that are not mathematically tractable. Computer simulation was once limited to mainframe computers. Currently, due to increased computer speeds, simulation studies can be completed on desktop and laptop computers in a reasonable period of time.

The steps to a simulation study are as follows.

1. Specify the method by which the test is to be carried out. This includes specifying how the test statistic is calculated and how the significance level is specified.
2. Generate a random sample,  $X_1, X_2, \dots, X_n$ , from the distribution specified by the alternative hypothesis. Calculate the test statistic from the simulated data and determine if the null hypothesis is accepted or rejected. Each of these samples is used to calculate the power of the test.
3. Generate a random sample,  $Y_1, Y_2, \dots, Y_n$ , from the distribution specified by the null hypothesis. Calculate the test statistic from the simulated data and determine if the null hypothesis is accepted or rejected. Each of these samples is used to calculate the significance level of the test.
4. Repeat steps 2 and 3 several thousand times, tabulating the number of times the simulated data lead to a rejection of the null hypothesis. The power is the proportion of simulation samples in step 2 that lead to rejection. The significance level is the proportion of simulated samples in step 3 that lead to rejection.

## Tests for One Mean (Simulation)

### Data Distributions

A wide variety of distributions may be studied. These distributions can vary in skewness, elongation, or other features such as bimodality. A detailed discussion of the distributions that may be used in the simulation is provided in the chapter 'Data Simulator'.

### 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 follows.

$$t_{n-1} = \frac{\bar{X} - M0}{s_{\bar{X}}/\sqrt{n}}$$

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  $M0$  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 (often 0.05), the null hypothesis is rejected. Otherwise, no conclusion can be reached.

#### 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 difference,  $M0$ , 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,  $W_R$ , is the minimum of  $Sp$  and  $Sn$ .
3. Compute the mean and standard deviation of  $W_R$  using the formulas

$$\mu_{W_R} = \frac{n(n+1)}{4}$$

$$\sigma_{W_R} = \sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum t^3 - \sum t}{48}}$$

where  $t$  represents the number of times the  $i^{\text{th}}$  value occurs.

4. Compute the z-value using

$$z_W = \frac{W_R - \mu_{W_R}}{\sigma_{W_R}}$$

## Tests for One Mean (Simulation)

For cases when  $n$  is less than 38, the significance level is found from a table of exact probabilities for the Wilcoxon test. When  $n$  is greater than or equal to 38, the significance of the test statistic is determined by comparing the  $z$  value to a normal probability table. If this  $p$ -value is less than a specified level (often 0.05), the null hypothesis is rejected. Otherwise, no conclusion can be reached.

### Sign Test

The sign test is popular because it is simple to compute. This test assumes that the data all follow the same distribution. The test is computed using the following steps.

1. Count the number of values strictly greater than  $M_0$ . Call this value  $X$ .
2. Count the number of values strictly less than  $M_0$ . Call this value  $Y$ .
3. Set  $m = X + Y$ .
4. Under the null hypothesis,  $X$  is distributed as a binomial random variable with a proportion of 0.5 and sample size of  $m$ .

The significance of  $X$  is calculated using binomial probabilities.

### Bootstrap Test

The one-sample bootstrap procedure for testing whether the mean is equal to a specific value is given in Efron & Tibshirani (1993), pages 224-227. The bootstrap procedure is as follows.

1. Compute the mean of the sample. Call it  $\bar{X}$ .
2. Compute the  $t$ -value using the standard  $t$ -test. The formula for this computation is

$$t_x = \frac{\bar{X} - M_0}{s_x / \sqrt{n}}$$

where  $M_0$  is the hypothesized mean.

3. Draw a random, with-replacement sample of size  $n$  from the original  $X$  values. Call this sample  $Y_1, Y_2, \dots, Y_n$ .
4. Compute the  $t$ -value of this bootstrap sample using the formula

$$t_y = \frac{\bar{Y} - \bar{X}}{s_y / \sqrt{n}}$$

5. For a two-tailed test, if  $|t_y| > |t_x|$  then add one to a counter variable,  $A$ .
6. Repeat steps 3 – 5  $B$  times.  $B$  may be anywhere from 100 to 10,000.
7. Compute the  $p$ -value of the bootstrap test as  $(A + 1) / (B + 1)$
8. Steps 1 – 7 complete one simulation iteration. Repeat these steps  $M$  times, where  $M$  is the number of simulations. The power and significance level are equal to the percent of the time the  $p$ -value is less than the nominal alpha of the test in their respective simulations.

Note that the bootstrap test is a time-consuming test to analyze, especially if you set  $B$  to a value much larger than 100.

## Tests for One Mean (Simulation)

### Exponential Test

The exponential distribution is a highly skewed distribution, so it is very different from the normal distribution. Thus, the t-test does not work well with exponential data.

There is an exact test for the mean of a sample drawn from the exponential distribution. It is well known that a simple function of the mean of exponential data follows the chi-square distribution. This relationship is given in Epstein (1960) as

$$\frac{2n\bar{X}}{M0} \sim \chi_{2n}^2$$

This expression can be used to test hypotheses about the value of the mean,  $M0$ .

---

### Standard Deviations

Care must be used when either the null or alternative distribution is not normal. In these cases, the standard deviation is usually not specified directly. For example, you might use a gamma distribution with a shape parameter of 1.5 and a mean of 4 as the null distribution and a gamma distribution with the same shape parameter and a mean of 5 as the alternative distribution. This allows you to compare the two means. However, note that although the shape parameters are constant, the standard deviations are not. In cases such as this, the null and alternatives not only have different means, but different standard deviations!

---

### 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 calculated using the values of the other parameters. Under most conditions, you would select either *Power* or *Sample Size*.

Select *Power* when you want to estimate the power for a specific scenario.

Select *Sample Size* when you want to determine the sample size needed to achieve a given power and alpha error level. This option can be very computationally intensive and may take considerable time to complete.

## Tests for One Mean (Simulation)

### Test and Simulations

#### Alternative Hypothesis

Specify the alternative hypothesis of the test. Since the null hypothesis is the opposite of the alternative, specifying the alternative is all that is needed. Usually, the two-tailed ( $\neq$ ) option is selected.

The options containing only  $<$  or  $>$  are one-tailed tests. When you choose one of these options, you must be sure that the input parameters match this selection.

Possible selections are:

- **Two-Sided (H1:  $\mu \neq \mu_0$ )**

This is the most common selection. It yields the *two-tailed t-test*. Use this option when you are testing whether the means are different, but you do not want to specify beforehand which mean is larger. Many scientific journals require two-tailed tests.

- **One-Sided (H1:  $\mu < \mu_0$ )**

This option yields a *one-tailed t-test*. Use it when you are only interested in the case in which  $\mu$  is less than  $\mu_0$ .

- **One-Sided (H1:  $\mu > \mu_0$ )**

This option yields a *one-tailed t-test*. Use it when you are only interested in the case in which  $\mu$  is greater than  $\mu_0$ .

#### Test Type

Specify which test statistic (t-test, Wilcoxon signed-rank test, sign test, bootstrap test, or exponential test) is to be simulated. Although the t-test is the most commonly used test statistic, it is based on assumptions that may not be viable in many situations. For your data, you may find that one of the other tests is more accurate (actual alpha = target alpha) and more precise (higher power).

Note that the bootstrap test is computationally intensive, so it can be very slow to evaluate.

#### Bootstrap Iterations

*Displayed when Test Type = Bootstrap Test*

Specify the number of iterations used in the bootstrap hypothesis test. This value is only used if the bootstrap test is displayed on the reports. The running time of the procedure depends heavily on the number of iterations specified here.

Recommendations by authors of books discussing the bootstrap range from 100 to 10,000. If you enter a large (greater than 500) value, the procedure may take several hours to run.

#### Simulations

This option specifies the number of iterations,  $M$ , used in the simulation. Larger numbers of iterations result in longer running time and more accurate results.

The precision of the simulated power estimates can be determined by recognizing that they follow the binomial distribution. Thus, confidence intervals may be constructed for power estimates. The following table gives an estimate of the precision that is achieved for various simulation sizes when the power is either 0.50 or 0.95. The table values are interpreted as follows: a 95% confidence interval of the true power is given by the power reported by the simulation plus and minus the 'Precision' amount given in the table.

### Tests for One Mean (Simulation)

Simulation Size M	Precision when Power = 0.50	Precision when Power = 0.95
100	0.100	0.044
500	0.045	0.019
1000	0.032	0.014
2000	0.022	0.010
5000	0.014	0.006
10000	0.010	0.004
50000	0.004	0.002
100000	0.003	0.001

Notice that a simulation size of 1000 gives a precision of plus or minus 0.014 when the true power is 0.95. Also note that as the simulation size is increased beyond 5000, there is only a small amount of additional precision achieved.

---

## 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. This value must be an integer greater than one. Note that you may enter a list of values using the syntax *50 100 150 200 250* or *50 to 250 by 50*.

## Tests for One Mean (Simulation)

---

### Effect Size

#### Input Type

Select which type of input method to use to specify the simulated distributions and the effect size.

The choices are

- **Simple**  
Provides the simplest possible input for specifying the distributions and effect size. This entry type is restricted to a set of distributions that can be specified using a mean and standard deviation. All simulated values come from one named distribution.
- **General**  
Provides all options for specifying the distributions and parameters. This allows for complete flexibility in specifying the simulated distributions but is more complicated.

---

### Effect Size – Simulated Distribution

*Displayed when Input Type = Simple*

#### Distribution to Simulate

Specify the distribution that represents the type of data you expect from your study.

The possible distributions are Beta, Exponential, Gamma, Gumbel, Laplace, Logistic, Lognormal, Normal, Poisson, TukeyGH, Uniform, Weibull.

All these distributions can be specified with a mean and standard deviation. The Beta and TukeyGH distributions each require two additional parameters. The Exponential and Poisson distributions require only the mean to be specified since the standard deviation can be computed from the mean.

#### Normal vs. TukeyGH

Tukey's distribution can be used to generate values that are nearly Normal, with departure from normality controlled by entering skewness (G) and kurtosis (H) parameters. Tukey's distribution with  $G = H = 0$  is equivalent to the Normal distribution.

---

### Effect Size – Means

*Displayed when Input Type = Simple*

#### $\mu_0$ (Null or Baseline Mean)

Enter a value for the population mean under the null hypothesis. This is the reference or baseline mean. If you are analyzing a paired  $t$ -test, this value should be zero.

Only the difference between  $\mu_1$  and  $\mu_0$  is used in the calculations.

#### $\mu_1$ (Actual Mean)

Enter a value (or range of values) for the actual population mean at which power and sample size are calculated.

Only the difference between  $\mu_1$  and  $\mu_0$  is used in the calculations.

## Tests for One Mean (Simulation)

---

**Effect Size – Standard Deviation***Displayed when Input Type = Simple* **$\sigma$  (Standard Deviation)**

This option specifies one or more values of the standard deviation. This must be a positive value. Be sure to use the standard deviation of  $X$  and not the standard deviation of the mean (the standard error). If you are doing a paired test, this is the standard deviation of the differences.

When this value is not known, you must supply an estimate of it. **PASS** includes a special tool for estimating the standard deviation. This tool may be loaded by pressing the *SD* button. Refer to the Standard Deviation Estimator chapter for further details.

---

**Effect Size – Simulated Distributions***Displayed when Input Type = General***Distribution|H0 (Null Hypothesis)**

Specify the distribution under the null hypothesis, H0.

The parameters of the distribution can be specified using numbers or letters. If letters are used, their values are specified in the boxes below. The value “M0” is reserved for the value of the mean under the null hypothesis.

Following is a list of the distributions that are available, and the syntax used to specify them. Each of the parameters should be replaced with a number or parameter name.

**Distributions with Common Parameters**

Beta(Shape1, Shape2, Min, Max)

Binomial(P, N)

Cauchy(Mean, Scale)

Constant(Value)

Exponential(Mean)

Gamma(Shape, Scale)

Gumbel(Location, Scale)

Laplace(Location, Scale)

Logistic(Location, Scale)

Lognormal(Mu, Sigma)

Multinomial(P1, P2, P3, ..., Pk)

Normal(Mean, Sigma)

Poisson(Mean)

TukeyGH(Mu, S, G, H)

Uniform(Min, Max)

Weibull(Shape, Scale)



## Tests for One Mean (Simulation)

### Distributions with Mean and SD Parameters

BetaMS(Mean, SD, Min, Max)

BinomialMS(Mean, N)

GammaMS(Mean, SD)

GumbelMS(Mean, SD)

LaplaceMS(Mean, SD)

LogisticMS(Mean, SD)

LognormalMS(Mean, SD)

UniformMS(Mean, SD)

WeibullMS(Mean, SD)

Details of writing mixture distributions, combined distributions, and compound distributions are found in the chapter on *Data Simulation* and will not be repeated here.

### Finding the Value of the Mean of a Specified Distribution

The mean of a distribution created as a linear combination of other distributions is found by applying the linear combination to the individual means. However, the mean of a distribution created by multiplying or dividing other distributions is not necessarily equal to applying the same function to the individual means. For example, the mean of  $4 \text{ Normal}(4, 5) + 2 \text{ Normal}(5, 6)$  is  $4*4 + 2*5 = 26$ , but the mean of  $4 \text{ Normal}(4, 5) * 2 \text{ Normal}(5, 6)$  is not exactly  $4*4*2*5 = 160$  (although it is close).

### Specifying the Mean for Paired (Matched) Data

Depending on the formula that is entered, the mean is not necessarily the value of  $M_0$ . For example, a common use of the one-group t-test is to test whether the mean of a set of differences is zero. Differences may be specified (ignoring the correlation between paired observations) as the difference between two normal distributions. This would be specified as  $\text{Normal}(M_0, S) - \text{Normal}(M_0, S)$ . The mean of the resulting distribution is  $M_0 - M_0 = 0$  (not  $M_0$ ).

### Distribution|H1

This option specifies the mean and distribution under the alternative hypothesis,  $H_1$ . That is, this is the actual (true) value of the mean at which the power is computed. Usually, the mean is specified by entering 'M1' for the mean parameter in the distribution expression and then entering values for the M1 parameter.

The parameters of the distribution can be specified using numbers or letters. If letters are used, their values are specified in the boxes below. The value "M1" is reserved for the value of the mean under the alternative hypothesis.

A list of the distributions that are available and the syntax used to specify them is given above.

### Finding the Value of the Mean under H1

The distributions have been parameterized in terms of their means, since this is the parameter being tested. The mean of a distribution created as a linear combination of other distributions is found by applying the linear combination to the individual means. However, the mean of a distribution created by multiplying or dividing other distributions is not necessarily equal to applying the same function to the individual means. For example, the mean of  $4 \text{ Normal}(4, 5) + 2 \text{ Normal}(5, 6)$  is  $4*4 + 2*5 = 26$ , but the mean of  $4 \text{ Normal}(4, 5) * 2 \text{ Normal}(5, 6)$  is not exactly  $4*4*2*5 = 160$  (although it is close).

## Tests for One Mean (Simulation)

### Specifying the Mean for Paired (Matched) Data

Depending on the formula that is entered, the mean is not necessarily the value of  $M1$ . For example, a common use of the one-group t-test is to test whether the mean of a set of differences is zero. Differences may be specified (ignoring the correlation between paired observations) as the difference between two normal distributions. This would be specified as  $\text{Normal}(M1, S) - \text{Normal}(M0, S)$ . The mean of the resulting distribution is  $M1 - M0$ .

---

### Effect Size – Distribution Parameters

*Displayed when Input Type = General*

#### **M0 (Mean|H0)**

These values are substituted for the  $M0$  in the distribution specifications given above.  $M0$  is intended to be the value of the mean hypothesized by the null hypothesis,  $H0$ .

You can enter a list of values using syntax such as *0 1 2 3* or *0 to 3 by 1*.

Note that whether  $M0$  is the mean of the simulated distribution depends on the formula you have entered. For example,  $\text{Normal}(M0, S)$  has a mean of  $M0$ , but  $\text{Normal}(M0, S) - \text{Normal}(M0, S)$  has a mean of zero.

#### **M1 (Mean|H1)**

These values are substituted for the  $M1$  in the distribution specifications given above.  $M1$  is intended to be the value of the mean hypothesized by the alternative hypothesis,  $H1$ .

You can enter a list of values using syntax such as *0 1 2 3* or *0 to 3 by 1*.

Note that whether  $M1$  is the mean of the simulated distribution depends on the formula you have entered. For example,  $\text{Normal}(M1, S)$  has a mean of  $M1$ , but  $\text{Normal}(M1, S) - \text{Normal}(M0, S)$  has a mean of  $M1 - M0$ .

#### **Parameter Values (S, A, B, C)**

Enter the numeric value(s) of parameter listed above. These values are substituted for the corresponding letter in the distribution specifications for  $H0$  and  $H1$ .

You can enter a list of values using syntax such as *0 1 2 3* or *0 to 3 by 1*.

You can also change the letter that is used as the name of this parameter.

## Tests for One Mean (Simulation)

## Example 1 – Power at Various Sample Sizes

A researcher is planning an experiment to test whether the mean response level to a certain drug is significantly different from zero. The researcher wants to use a t-test with an alpha level of 0.05. He wants to compute the power at various sample sizes from 5 to 40, assuming the true mean is one. He assumes that the data are normally distributed with a standard deviation of 2.

### 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 One Mean (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 1a** by going to the **File** menu and choosing **Open Example Template**. To analyze the same problem using the General input type, load the Example 1b settings template.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: <math>\mu \neq \mu_0</math>)</b>
Test Type .....	<b>T-Test</b>
Simulations .....	<b>2000</b>
Alpha .....	<b>0.05</b>
N (Sample Size) .....	<b>5 to 40 by 5</b>
Input Type .....	<b>Simple</b>
Distribution to Simulate .....	<b>Normal</b>
$\mu_0$ (Null or Baseline Mean) .....	<b>0</b>
$\mu_1$ (Actual Mean) .....	<b>1</b>
$\sigma$ (Standard Deviation) .....	<b>2</b>

### Annotated Output

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

### Numeric Results

<b>Numeric Results</b>									
Hypotheses: $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$									
Test Statistic: T-Test									
Simulated Distribution: Normal									
Power	N	$\mu_0$	$\mu_1$	Diff $\mu_1 - \mu_0$	$\sigma$	Effect Size	Target Alpha	Actual Alpha	
0.1415 (0.0153) [0.1262 0.1568]	5	0.0	1.0	1.0	2.0	0.500	0.050	0.056 (0.010) [0.046 0.066]	
0.2910 (0.0199) [0.2711 0.3109]	10	0.0	1.0	1.0	2.0	0.500	0.050	0.051 (0.010) [0.041 0.061]	
0.4330 (0.0217) [0.4113 0.4547]	15	0.0	1.0	1.0	2.0	0.500	0.050	0.051 (0.010) [0.041 0.061]	
0.5620 (0.0217) [0.5403 0.5837]	20	0.0	1.0	1.0	2.0	0.500	0.050	0.049 (0.009) [0.040 0.058]	
0.6660 (0.0207) [0.6453 0.6867]	25	0.0	1.0	1.0	2.0	0.500	0.050	0.042 (0.009) [0.033 0.050]	

Tests for One Mean (Simulation)

0.7595 (0.0187)	30 [0.7408 0.7782]	0.0	1.0	1.0	2.0	0.500	0.050	0.050 (0.010) [0.040 0.060]
0.8200 (0.0168)	35 [0.8032 0.8368]	0.0	1.0	1.0	2.0	0.500	0.050	0.049 (0.009) [0.040 0.058]
0.8840 (0.0140)	40 [0.8700 0.8980]	0.0	1.0	1.0	2.0	0.500	0.050	0.053 (0.010) [0.043 0.062]

Notes:

Simulations: 2000. Run Time: 3.25 seconds.

Report Definitions

Power is the probability of rejecting a false null hypothesis. Second Row: (Power Inc.) [95% LCL and UCL Power].

N is the size of the sample drawn from the population.

$\mu_0$  is the value of the population mean under the null hypothesis.

$\mu_1$  is the actual value of the population mean at which power and sample size are calculated.

$\mu_1 - \mu_0$  is the difference between the actual and null means.

$\sigma$  is the standard deviation of the population. It measures the variability in the population.

Effect Size =  $|\mu_1 - \mu_0|/\sigma$  is the relative magnitude of the effect.

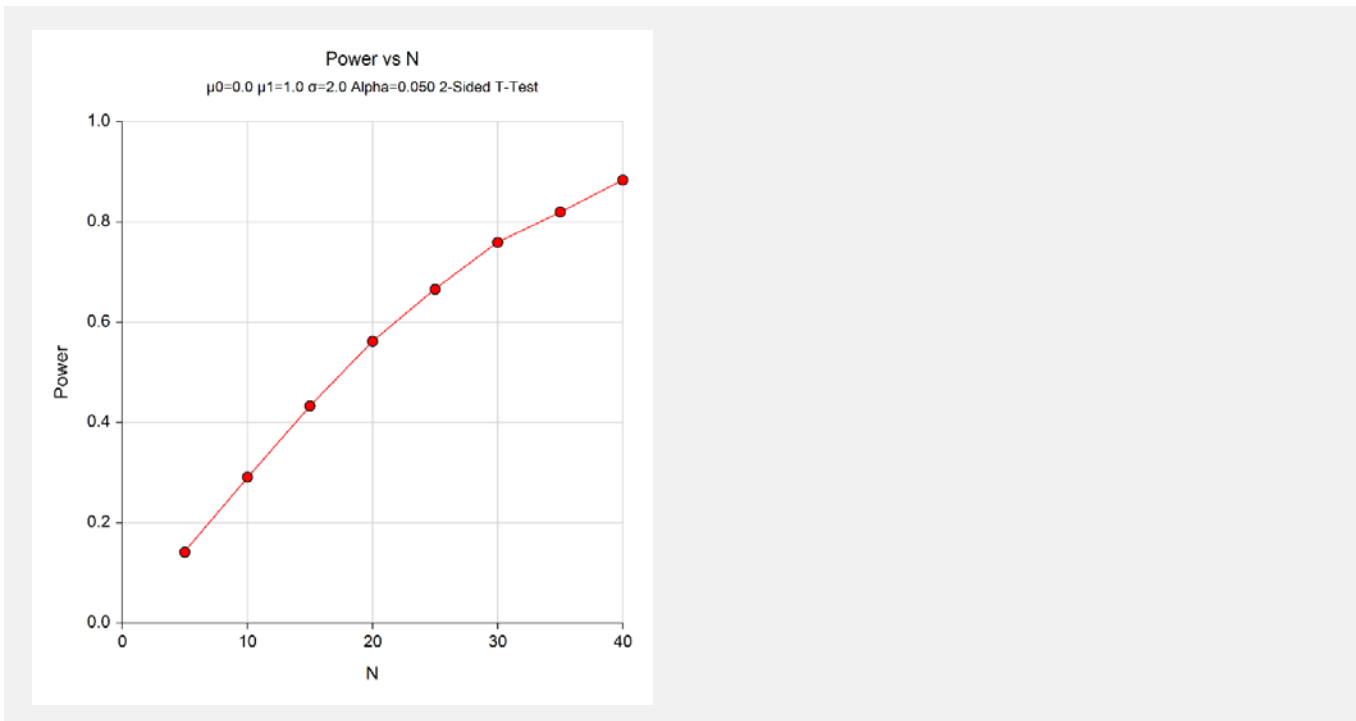
Target Alpha is the probability of rejecting a true null hypothesis. It is set by the user.

Actual Alpha is the alpha level that was actually achieved by the experiment. Second Row: (Alpha Inc.) [95% LCL and UCL Alpha].

This report shows the estimated power for each scenario. The first row shows the parameter settings and the estimated power and significance level (Actual Alpha). Note that because these are results of a simulation study, the computed power and alpha will vary from run to run. Thus, another report obtained using the same input parameters will be slightly different than the one above.

The second row shows two 95% confidence intervals in brackets: the first for the power and the second for the significance level. Half the width of each confidence interval is given in parentheses as a fundamental measure of the accuracy of the simulation. As the number of simulations is increased, the width of the confidence interval will decrease.

Plots Section



This plot shows the relationship between sample size and power.

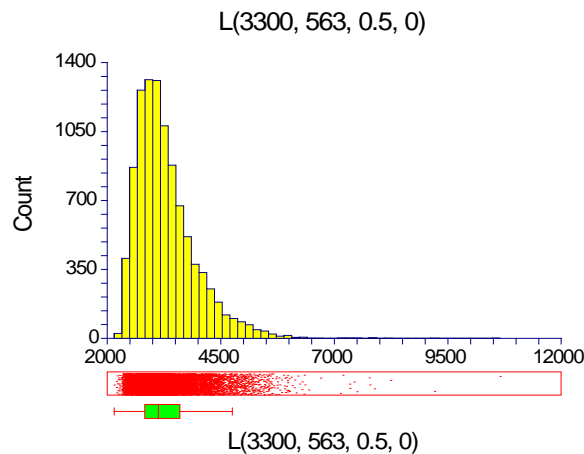
## Tests for One Mean (Simulation)

## Example 2 – Finding the Sample Size for Skewed Data

In studying deaths from SIDS (Sudden Infant Death Syndrome), one hypothesis put forward is that infants dying of SIDS weigh less than normal at birth. Suppose the average birth weight of infants is 3300 grams with a standard deviation of 663 grams. The researchers decide to examine the effect of a skewed distribution on the test used by adding skewness to the simulated data using Tukey's Lambda distribution with a skewness factor of 0.5.

Using the Data Simulator program, the researchers found that the actual standard deviation using the above parameters was almost 800. This occurs because adding skewness changes the standard deviation. They found that setting the standard deviation in Tukey's Lambda distribution to 563 resulted in a standard deviation in the data of about 663.

A histogram of 10,000 pseudo-random values from this distribution appears as follows.



The researchers want to determine how large a sample of SIDS infants will be needed to detect a drop in average weight of 25%? Note that applying this percentage to the average weight of 3300 yields 2475. Use an alpha of 0.05 and 80% power.

Although a one-sided hypothesis might be considered, sample size estimates will assume a two-sided alternative to keep the research design in line with other studies.

## 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 One Mean (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 2a** by going to the **File** menu and choosing **Open Example Template**. To analyze the same problem using the General input type, load the Example 2b settings template.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: <math>\mu \neq \mu_0</math>)</b>
Test Type .....	<b>T-Test</b>
Simulations .....	<b>2000</b>
Power .....	<b>0.80</b>
Alpha .....	<b>0.05</b>
Input Type .....	<b>Simple</b>
Distribution to Simulate .....	<b>TukeyGH</b>
G (Skewness) .....	<b>0.5</b>
H (Kurtosis) .....	<b>0</b>

## Tests for One Mean (Simulation)

## Design Tab (Continued)

$\mu_0$ (Null or Baseline Mean) .....	2475
$\mu_1$ (Actual Mean) .....	3300
$\sigma$ (Standard Deviation) .....	563

## Output

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

## Numeric Results of Search for N

Numeric Results										
Hypotheses: $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$										
Test Statistic: T-Test										
Simulated Distribution: TukeyGH										
Power	N	$\mu_0$	$\mu_1$	Diff $\mu_1 - \mu_0$	$\sigma$	Effect Size	Target Alpha	Actual Alpha	G	H
0.8305 (0.0164)	5	2475.0	3300.0	825.0	563.0	1.465	0.050	0.084 (0.012)	0.5 [0.071 0.096]	0.0

The required sample size was 5. Note that the actual alpha value is between 0.071 and 0.096, which is greater than 0.05. This shows one of the problems of using the t-test with a skewed distribution.

To be more accurate and yet avoid the long running time of the search for N, a reasonable strategy would be to run simulations to obtain the powers using N's from 4 to 10. The result of this study is displayed next (Example 2c).

## Numeric Results of Power Search for Various N

Numeric Results										
Hypotheses: $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$										
Test Statistic: T-Test										
Simulated Distribution: TukeyGH										
Power	N	$\mu_0$	$\mu_1$	Diff $\mu_1 - \mu_0$	$\sigma$	Effect Size	Target Alpha	Actual Alpha	G	H
0.6320 (0.0211)	4	2475.0	3300.0	825.0	563.0	1.465	0.050	0.078 (0.012)	0.5 [0.066 0.089]	0.0
0.8290 (0.0165)	5	2475.0	3300.0	825.0	563.0	1.465	0.050	0.083 (0.012)	0.5 [0.071 0.095]	0.0
0.9455 (0.0099)	6	2475.0	3300.0	825.0	563.0	1.465	0.050	0.070 (0.011)	0.5 [0.059 0.081]	0.0
0.9805 (0.0061)	7	2475.0	3300.0	825.0	563.0	1.465	0.050	0.092 (0.013)	0.5 [0.079 0.104]	0.0
0.9955 (0.0029)	8	2475.0	3300.0	825.0	563.0	1.465	0.050	0.089 (0.012)	0.5 [0.076 0.101]	0.0
0.9960 (0.0028)	9	2475.0	3300.0	825.0	563.0	1.465	0.050	0.078 (0.012)	0.5 [0.066 0.089]	0.0
0.9980 (0.0020)	10	2475.0	3300.0	825.0	563.0	1.465	0.050	0.067 (0.011)	0.5 [0.056 0.077]	0.0

The sample sizes of 6 or 10 appears to meet the design parameters the best, but the actual significance level still appears to be greater than 0.05. The researchers decide that they must use a smaller value of Alpha so that the actual alpha is about 0.05. After some experimentation, they find that setting Alpha to 0.025 results in the desired power and significance level (Example 2d).

## Tests for One Mean (Simulation)

## Numeric Results with Alpha = 0.025

## Numeric Results

Hypotheses:  $H_0: \mu = \mu_0$  vs.  $H_1: \mu \neq \mu_0$ 

Test Statistic: T-Test

Simulated Distribution: TukeyGH

Power	N	$\mu_0$	$\mu_1$	Diff $\mu_1 - \mu_0$	$\sigma$	Effect Size	Target Alpha	Actual Alpha	G	H
0.3570 (0.0210) [0.3360 0.3780]	4	2475.0	3300.0	825.0	563.0	1.465	0.025	0.045 (0.009) [0.036 0.054]	0.5	0.0
0.6195 (0.0213) [0.5982 0.6408]	5	2475.0	3300.0	825.0	563.0	1.465	0.025	0.050 (0.010) [0.040 0.060]	0.5	0.0
<b>0.8175</b> <b>(0.0169) [0.8006 0.8344]</b>	<b>6</b>	<b>2475.0</b>	<b>3300.0</b>	<b>825.0</b>	<b>563.0</b>	<b>1.465</b>	<b>0.025</b>	<b>0.041</b> <b>(0.009) [0.032 0.049]</b>	<b>0.5</b>	<b>0.0</b>
0.9045 (0.0129) [0.8916 0.9174]	7	2475.0	3300.0	825.0	563.0	1.465	0.025	0.042 (0.009) [0.033 0.050]	0.5	0.0
0.9660 (0.0079) [0.9581 0.9739]	8	2475.0	3300.0	825.0	563.0	1.465	0.025	0.056 (0.010) [0.045 0.066]	0.5	0.0
0.9805 (0.0061) [0.9744 0.9866]	9	2475.0	3300.0	825.0	563.0	1.465	0.025	0.047 (0.009) [0.038 0.056]	0.5	0.0
0.9895 (0.0045) [0.9850 0.9940]	10	2475.0	3300.0	825.0	563.0	1.465	0.025	0.037 (0.008) [0.029 0.045]	0.5	0.0

It appears that a sample size of 6 with a Target Alpha of 0.025 will result in an experimental design with the characteristics the researchers wanted.

Notice that when working with non-normal distributions, you must change both N and the Target Alpha to achieve the design you want!

## Example 3 – Comparative Results with Skewed Data

Continuing with Example 2, the researchers want to study the characteristics of various test statistics as the amount of skewness is increased. To do this, they let the skewness parameter of Tukey's Lambda distribution vary between 0 and 1. The researchers realize that the standard deviation will change as the skewness parameter is increased, but they decide to ignore this complication.

### 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 One Mean (Simulation)** procedure window. 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**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: <math>\mu \neq \mu_0</math>)</b>
Test Type .....	<b>T-Test</b>
Simulations .....	<b>2000</b>
Alpha .....	<b>0.05</b>
N (Sample Size) .....	<b>6</b>
Input Type .....	<b>Simple</b>
Distribution to Simulate .....	<b>TukeyGH</b>
G (Skewness) .....	<b>0.0 0.2 0.4 0.6 0.8 1.0</b>
H (Kurtosis) .....	<b>0</b>
$\mu_0$ (Null or Baseline Mean) .....	<b>2475</b>
$\mu_1$ (Actual Mean) .....	<b>3300</b>
$\sigma$ (Standard Deviation) .....	<b>563</b>
<b>Reports Tab</b>	
Show Comparative Reports .....	<b>Checked</b>
Include T-Test Results .....	<b>Checked</b>
Include Wilcoxon & Sign Test .....	<b>Checked</b>
<b>Comparative Plots Tab</b>	
Show Comparative Plots .....	<b>Checked</b>



Tests for One Mean (Simulation)

Output

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

Numeric Results

Power Comparison

Hypotheses:  $H_0: \mu = \mu_0$  vs.  $H_1: \mu \neq \mu_0$   
 Simulated Distribution: TukeyGH

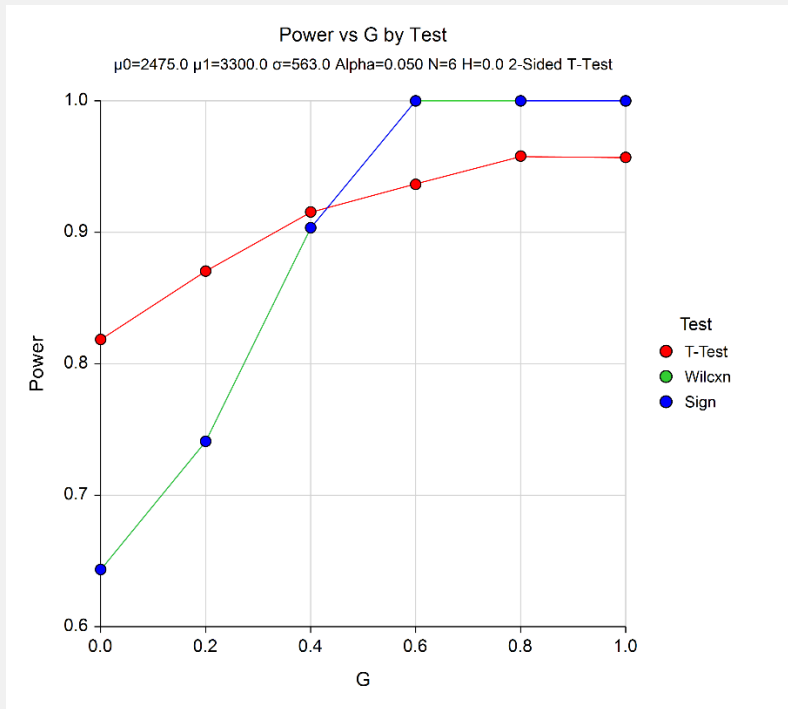
N	$\mu_0$	$\mu_1$	$\sigma$	Target Alpha	T-Test Power	Wilcxn Power	Sign Power	G	H
6	2475.0	3300.0	563.0	0.050	0.8175	0.6340	0.6340	0.0	0.0
6	2475.0	3300.0	563.0	0.050	0.8805	0.7485	0.7485	0.2	0.0
6	2475.0	3300.0	563.0	0.050	0.9350	0.9025	0.9025	0.4	0.0
6	2475.0	3300.0	563.0	0.050	0.9455	1.0000	1.0000	0.6	0.0
6	2475.0	3300.0	563.0	0.050	0.9455	1.0000	1.0000	0.8	0.0
6	2475.0	3300.0	563.0	0.050	0.9575	1.0000	1.0000	1.0	0.0

Alpha Comparison

Hypotheses:  $H_0: \mu = \mu_0$  vs.  $H_1: \mu \neq \mu_0$   
 Simulated Distribution: TukeyGH

N	$\mu_0$	$\mu_1$	$\sigma$	Target Alpha	T-Test Alpha	Wilcxn Alpha	Sign Alpha	G	H
6	2475.0	3300.0	563.0	0.050	0.049	0.027	0.027	0.0	0.0
6	2475.0	3300.0	563.0	0.050	0.067	0.043	0.043	0.2	0.0
6	2475.0	3300.0	563.0	0.050	0.069	0.041	0.041	0.4	0.0
6	2475.0	3300.0	563.0	0.050	0.102	0.061	0.061	0.6	0.0
6	2475.0	3300.0	563.0	0.050	0.140	0.088	0.088	0.8	0.0
6	2475.0	3300.0	563.0	0.050	0.182	0.113	0.113	1.0	0.0

Chart Section



Several interesting trends become apparent from this study. First, for a sample size of 6, the power of the Wilcoxon test and the sign test are the same (this is not the case for larger sample sizes). The alpha of the t-test increases as the amount of skewness increases. The alpha of the Wilcoxon and sign tests does not increase as rapidly as it does for the t-test.

## Tests for One Mean (Simulation)

**Example 4 – Validation using Zar (1984)**

Zar (1984), pages 111-112, presents an example in which  $\mu_0 = 0.0$ ,  $\mu_1 = 1.0$ ,  $\sigma = 1.25$ ,  $\alpha = 0.05$ , and  $N = 12$ . Zar obtains an approximate power of 0.72. We will validate this procedure by running this example. To make certain that the results are very accurate, the number of simulations will be set to 10,000.

**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 One Mean (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 4** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: <math>\mu \neq \mu_0</math>)</b>
Test Type .....	<b>T-Test</b>
Simulations .....	<b>10000</b>
Alpha .....	<b>0.05</b>
N (Sample Size) .....	<b>12</b>
Input Type .....	<b>Simple</b>
Distribution to Simulate .....	<b>Normal</b>
$\mu_0$ (Null or Baseline Mean) .....	<b>0</b>
$\mu_1$ (Actual Mean) .....	<b>1</b>
$\sigma$ (Standard Deviation) .....	<b>1.25</b>

**Output**

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

**Numeric Results**

<b>Numeric Results</b>									
Hypotheses: $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$									
Test Statistic: T-Test									
Simulated Distribution: Normal									
Power	N	$\mu_0$	$\mu_1$	Diff $\mu_1 - \mu_0$	$\sigma$	Effect Size	Target Alpha	Actual Alpha	
0.7168	12	0.0	1.0	1.0	1.3	0.800	0.050	0.048	
(0.0088)	[0.7080	0.7256]						(0.004)	[0.044 0.053]

This simulation obtained a power of 0.7168 which rounds to the 0.72 computed by Zar. Note that another repetition of this same analysis will probably be slightly different since a different set of random numbers will be used.

## Tests for One Mean (Simulation)

**Example 5 – Validation using Machin (1997)**

Machin, et. al. (1997), page 37, present an example in which  $\mu_0 = 0.0$ ,  $\mu_1 = 0.2$ ,  $\sigma = 1.0$ ,  $\alpha = 0.05$ , and power = 0.80. They obtain a sample size of 199.

**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 One Mean (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 5** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: <math>\mu \neq \mu_0</math>)</b>
Test Type .....	<b>T-Test</b>
Simulations .....	<b>2000</b>
Power .....	<b>0.80</b>
Alpha .....	<b>0.05</b>
Input Type .....	<b>Simple</b>
Distribution to Simulate .....	<b>Normal</b>
$\mu_0$ (Null or Baseline Mean) .....	<b>0</b>
$\mu_1$ (Actual Mean) .....	<b>0.2</b>
$\sigma$ (Standard Deviation) .....	<b>1</b>

**Output**

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

**Numeric Results**

<b>Numeric Results</b>									
Hypotheses: $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$									
Test Statistic: T-Test									
Simulated Distribution: Normal									
Power	N	$\mu_0$	$\mu_1$	Diff $\mu_1 - \mu_0$	$\sigma$	Effect Size	Target Alpha	Actual Alpha	
0.8125 (0.0171)	<b>199</b> [0.7954 0.8296]	0.0	0.2	0.2	1.0	0.200	0.050	0.063 (0.011)	[0.052 0.073]

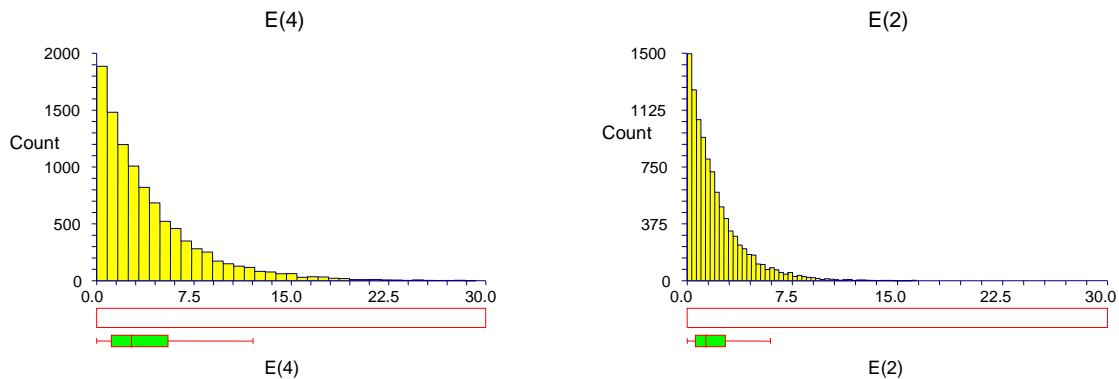
The sample size result matches the value of Machin (1997). If you run the simulation multiple times, you'll come up with values right around 199.

## Tests for One Mean (Simulation)

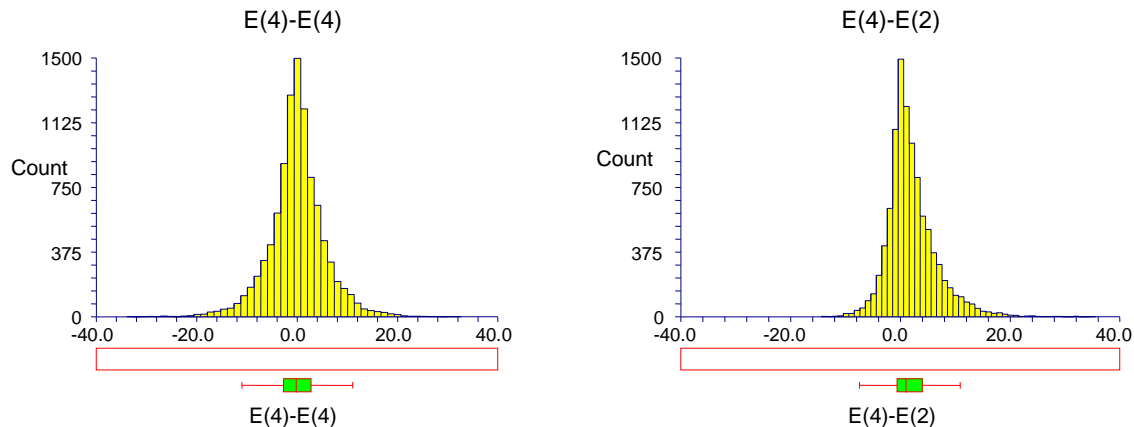
## Example 6 – Power of the Wilcoxon Signed-Rank Test

The Wilcoxon nonparametric signed-rank test was designed for data that do not follow the normal distribution but are symmetric. This type of data often occurs when differences between two non-normal variables are taken, as in a study that analyzes differences in pre- and post-test scores.

For this example, suppose the pre-test and the post-test scores are exponentially distributed. Here are examples of exponentially-distributed data with means of 4 and 2, respectively.



It has been shown that the differences between two identically-distributed variables are symmetric. The histogram below on the left shows differences in the null case in which the difference is between two exponential variables both with a mean of 4. The histogram below on the right shows differences in the alternative case in which the difference is between an exponential variable with a mean of 4 and an exponential variable with a mean of 2. Careful inspection shows that the second histogram is skewed to the right and the mean difference is about 2, not 0.



The researchers want to study the power of the two-sided Wilcoxon signed-rank test when sample sizes of 10, 20, 30, and 40 are used, and testing is done at the 5% significance level.

## Tests for One Mean (Simulation)

## 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 One Mean (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 6** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: <math>\mu \neq \mu_0</math>)</b>
Test Type .....	<b>Wilcoxon Signed-Rank Test</b>
Simulations .....	<b>5000</b>
Alpha .....	<b>0.05</b>
N (Sample Size) .....	<b>10 20 30 40 50</b>
Input Type .....	<b>General</b>
Distribution H0 .....	<b>Exponential(M0)-Exponential(M0)</b>
Distribution H1 .....	<b>Exponential(M0)-Exponential(M1)</b>
M0 (Mean H0) Parameter Value(s) .....	<b>4</b>
M1 (Mean H1) Parameter Value(s) .....	<b>2</b>

## Output

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

### Numeric Results

<b>Numeric Results</b>									
Hypotheses: H0: $\mu = \mu_0$ vs. H1: $\mu \neq \mu_0$									
Test Statistic: Wilcoxon Signed-Rank Test									
H0 Distribution: Exponential(M0)-Exponential(M0)									
H1 Distribution: Exponential(M0)-Exponential(M1)									
Power	N	$\mu_0$	$\mu_1$	Diff $\mu_1 - \mu_0$	Target Alpha	Actual Alpha	M0	M1	
0.2286 (0.0116)	10	0.0	2.0	2.0	0.050	0.048 (0.006)	4.0	2.0	[0.2170 0.2402] [0.042 0.054]
0.4548 (0.0138)	20	0.0	2.0	2.0	0.050	0.046 (0.006)	4.0	2.0	[0.4410 0.4686] [0.040 0.051]
0.6570 (0.0132)	30	0.0	2.0	2.0	0.050	0.049 (0.006)	4.0	2.0	[0.6438 0.6702] [0.043 0.055]
0.7744 (0.0116)	40	0.0	2.0	2.0	0.050	0.053 (0.006)	4.0	2.0	[0.7628 0.7860] [0.047 0.059]
0.8720 (0.0093)	50	0.0	2.0	2.0	0.050	0.050 (0.006)	4.0	2.0	[0.8627 0.8813] [0.044 0.056]

Reasonable power is achieved with N = 50.

## Tests for One Mean (Simulation)

**Example 7 – Likert-Scale Data**

Likert-scale data occurs commonly in survey research. A *Likert Scale* is discrete, ordinal data. It usually occurs when a survey poses a question and the respondent must pick among strongly agree, agree, undecided, disagree, or strongly disagree. The responses are usually coded as 1, 2, 3, 4, and 5.

Likert data can be analyzed in a number of ways. Perhaps the most common is to use a t-test or a Wilcoxon test. (Using the Wilcoxon test is invalid in this case because the data are seldom distributed symmetrically.)

In this example, a questionnaire is planned on which Likert-scale questions will be asked. The researchers want to study the power and actual significance levels of various sample sizes. They decide to look at what happens as the proportion of strongly agree responses is increased beyond a perfectly uniform response pattern. They want to compute the power when the strongly agree response is twice as likely, four times as likely, and eight times as likely. The sample size is 20, alpha is 0.05, and the test is two-sided.

**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 One Mean (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 7** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: <math>\mu \neq \mu_0</math>)</b>
Test Type .....	<b>T-Test</b>
Simulations .....	<b>5000</b>
Alpha .....	<b>0.05</b>
N (Sample Size) .....	<b>20</b>
Input Type .....	<b>General</b>
Distribution H0 .....	<b>Multinomial(M0 1 1 1 1)</b>
Distribution H1 .....	<b>Multinomial(M1 1 1 1 1)</b>
M0 (Mean H0) Parameter Value(s) .....	<b>1</b>
M1 (Mean H1) Parameter Value(s) .....	<b>2 4 8</b>

## Tests for One Mean (Simulation)

## Output

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

## Numeric Results

## Numeric Results

Hypotheses:  $H_0: \mu = \mu_0$  vs.  $H_1: \mu \neq \mu_0$

Test Statistic: T-Test

$H_0$  Distribution: Multinomial( $M_0$  1 1 1 1)

$H_1$  Distribution: Multinomial( $M_1$  1 1 1 1)

Power	N	$\mu_0$	$\mu_1$	Diff $\mu_1 - \mu_0$	Target Alpha	Actual Alpha	$M_0$	$M_1$
0.1692 (0.0104) [0.1588 0.1796]	20	3.0	2.7	-0.3	0.050	0.051 (0.006) [0.045 0.057]	1.0	2.0
0.5658 (0.0137) [0.5521 0.5795]	20	3.0	2.3	-0.8	0.050	0.057 (0.006) [0.051 0.064]	1.0	4.0
0.9162 (0.0077) [0.9085 0.9239]	20	3.0	1.8	-1.2	0.050	0.044 (0.006) [0.039 0.050]	1.0	8.0

Note that  $M_0$  and  $M_1$  are no longer the  $H_0$  and  $H_1$  means. Now, they represent the relative weighting given to the strongly agree response. Under  $H_0$ , the mean ( $\mu_0$ ) is 3.0. As  $M_1$  is increased, the mean under  $H_1$  ( $\mu_1$ ) changes from 2.7 to 2.3 to 1.8. We note that the actual significance level, alpha, remains close to the target value of 0.05.

## Example 8 – Computing the Power after Completing an Experiment

A group of researchers has completed an experiment designed to determine if a particular hormone increases weight gain in rats. The researchers inject 20 rats of the same age with the hormone and measure their weight gain after 1 month. The investigators uses the two-sided bootstrap test with  $\alpha = 0.05$  and 100 bootstrap samples to determine if the average weight gained by these rats (171 grams) is significantly greater than the known average weight gained by rats of the same age over the same period of time (155 grams). Unfortunately, the results indicate that there is no significant difference between the two means. Therefore, the researchers decide to compute the power achieved by this test for alternative means ranging from 160 to 190 grams. They decide to use 1000 simulations for the study. For comparative purposes, they also decide to look at the power achieved by the bootstrap test in comparison to various other applicable tests. Suppose that they know that the standard deviation for weight gain is 33 grams.

Note that the researchers compute the power for a range of practically significant alternatives. The range chosen should represent likely values based on historical evidence.

### 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 One Mean (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 8** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided (<math>H_1: \mu \neq \mu_0</math>)</b>
Test Type .....	<b>Bootstrap Test</b>
Bootstrap Iterations.....	<b>100</b>
Simulations .....	<b>2000</b>
Alpha .....	<b>0.05</b>
N (Sample Size).....	<b>20</b>
Input Type .....	<b>Simple</b>
Distribution to Simulate.....	<b>Normal</b>
$\mu_0$ (Null or Baseline Mean) .....	<b>155</b>
$\mu_1$ (Actual Mean) .....	<b>160 to 190 by 10</b>
$\sigma$ (Standard Deviation) .....	<b>33</b>
<b>Reports Tab</b>	
Show Comparative Reports.....	<b>Checked</b>
Include T-Test Results.....	<b>Checked</b>
Include Wilcoxon & Sign Test Results....	<b>Checked</b>
Include Bootstrap Test Results.....	<b>Checked</b>
<b>Comparative Plots Tab</b>	
Show Comparative Plots .....	<b>Checked</b>



## Tests for One Mean (Simulation)

## Output

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

## Numeric Results for Power of Bootstrap

## Numeric Results

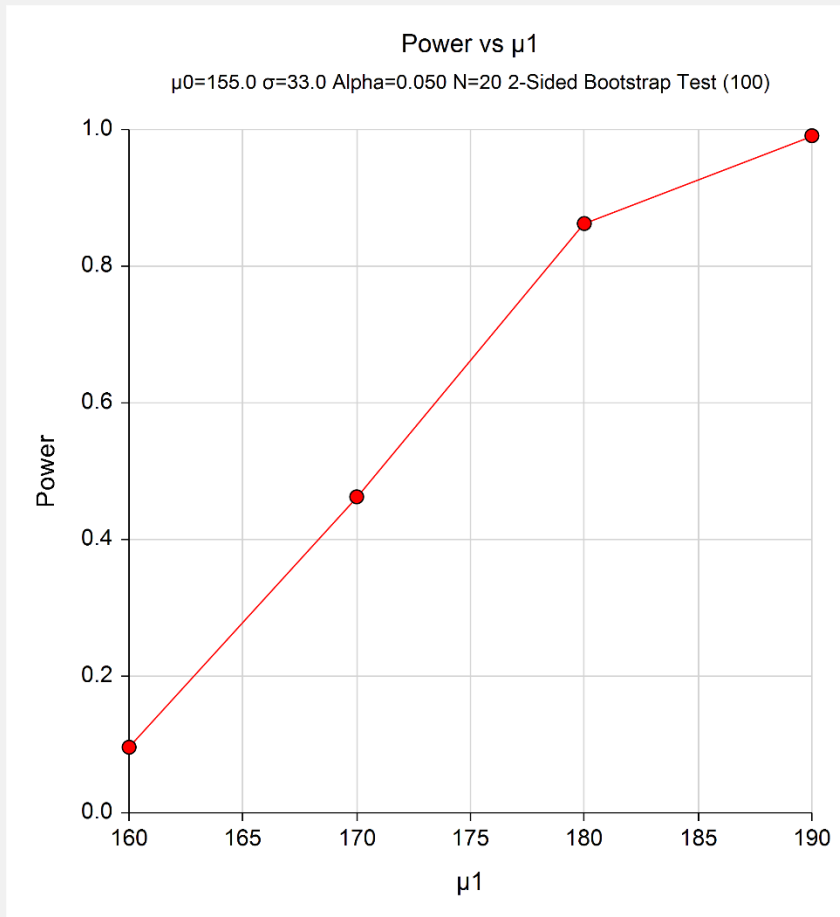
Hypotheses:  $H_0: \mu = \mu_0$  vs.  $H_1: \mu \neq \mu_0$

Test Statistic: Bootstrap Test (100)

Simulated Distribution: Normal

Power	N	$\mu_0$	$\mu_1$	Diff $\mu_1 - \mu_0$	$\sigma$	Effect Size	Target Alpha	Actual Alpha
0.0960 (0.0129) [0.0831 0.1089]	20	155.0	160.0	5.0	33.0	0.152	0.050	0.039 (0.008) [0.031 0.047]
0.4625 (0.0219) [0.4406 0.4844]	20	155.0	170.0	15.0	33.0	0.455	0.050	0.051 (0.010) [0.041 0.061]
0.8625 (0.0151) [0.8474 0.8776]	20	155.0	180.0	25.0	33.0	0.758	0.050	0.051 (0.010) [0.041 0.061]
0.9910 (0.0041) [0.9869 0.9951]	20	155.0	190.0	35.0	33.0	1.061	0.050	0.049 (0.009) [0.040 0.058]

## Chart Section



Reasonable power is achieved by this test for alternative means larger than 180. The accuracy of these results, of course, depends on the assumption that the data are normally distributed.

Tests for One Mean (Simulation)

Comparative Results for Power of Various Tests

Power Comparison

Hypotheses:  $H_0: \mu = \mu_0$  vs.  $H_1: \mu \neq \mu_0$   
 Simulated Distribution: Normal

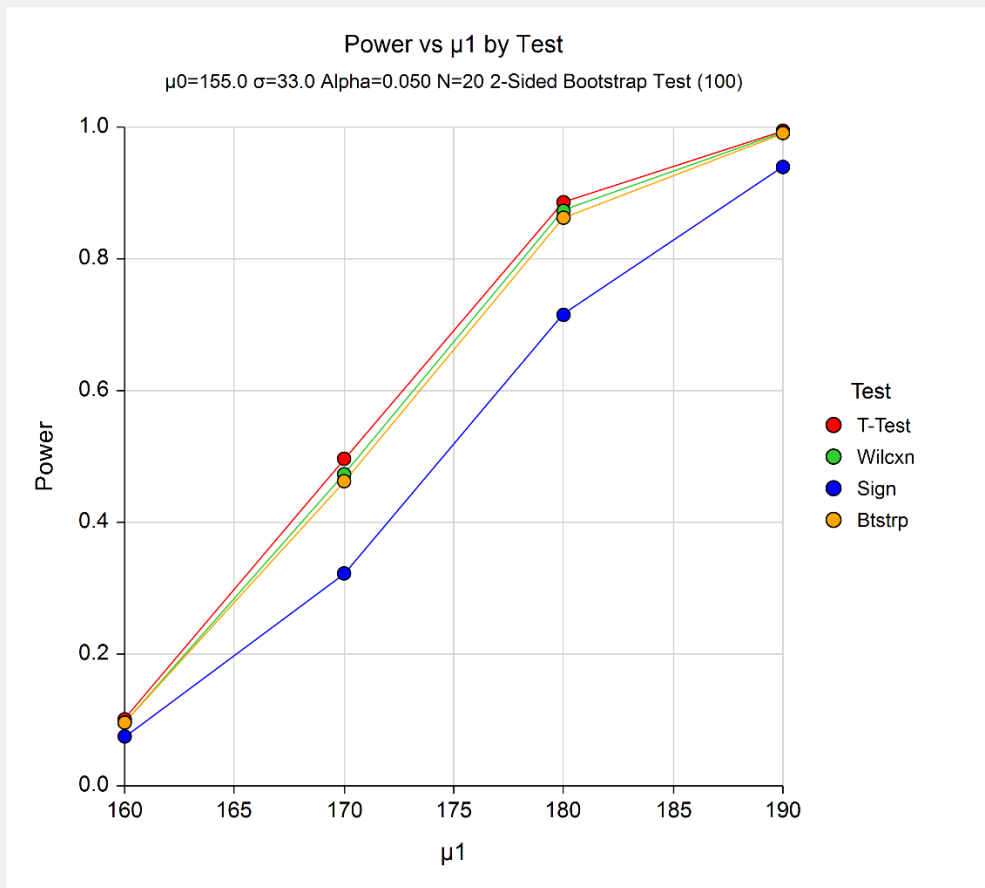
N	$\mu_0$	$\mu_1$	$\sigma$	Target Alpha	T-Test Power	Wilcxn Power	Sign Power	Btstrp Power
20	155.0	160.0	33.0	0.050	0.1010	0.0960	0.0750	0.0960
20	155.0	170.0	33.0	0.050	0.4965	0.4730	0.3225	0.4625
20	155.0	180.0	33.0	0.050	0.8860	0.8730	0.7150	0.8625
20	155.0	190.0	33.0	0.050	0.9945	0.9915	0.9395	0.9910

Alpha Comparison

Hypotheses:  $H_0: \mu = \mu_0$  vs.  $H_1: \mu \neq \mu_0$   
 Simulated Distribution: Normal

N	$\mu_0$	$\mu_1$	$\sigma$	Target Alpha	T-Test Alpha	Wilcxn Alpha	Sign Alpha	Btstrp Alpha
20	155.0	160.0	33.0	0.050	0.043	0.044	0.034	0.039
20	155.0	170.0	33.0	0.050	0.058	0.061	0.051	0.051
20	155.0	180.0	33.0	0.050	0.052	0.046	0.036	0.051
20	155.0	190.0	33.0	0.050	0.053	0.050	0.039	0.049

Chart Section



It is apparent from these results that the bootstrap performs about as well as the t-test and nonparametric tests for this design.

## Example 9 – Comparison of Tests for Exponential Data

A researcher is designing an experiment. She believes that the data will follow an exponential distribution. Consequently, she does not believe that the t-test will be useful for her situation. She would like to compare several possible tests to determine which would be best for analyzing exponential data. She is interested in determining the power when the alternative mean is twice the null mean, which is 10. She wants to find the power achieved for sample sizes ranging from 20 to 60 with  $\alpha = 0.05$ .

### 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 One Mean (Simulation)** procedure window. You may then make the appropriate entries as listed below, or open **Example 9** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided (<math>H_1: \mu \neq \mu_0</math>)</b>
Test Type .....	<b>T-Test</b>
Simulations .....	<b>2000</b>
Alpha .....	<b>0.05</b>
N (Sample Size) .....	<b>20 to 60 by 20</b>
Input Type .....	<b>Simple</b>
Distribution to Simulate .....	<b>Exponential</b>
$\mu_0$ (Null or Baseline Mean) .....	<b>10</b>
$\mu_1$ (Actual Mean) .....	<b>20</b>
<b>Reports Tab</b>	
Show Comparative Reports .....	<b>Checked</b>
Include T-Test Results .....	<b>Checked</b>
Include Wilcoxon & Sign Test Results ....	<b>Checked</b>
Include Bootstrap Test Results .....	<b>Checked</b>
Include Exponential Test Results .....	<b>Checked</b>
<b>Comparative Plots Tab</b>	
Show Comparative Plots .....	<b>Checked</b>

Tests for One Mean (Simulation)

Output

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

Numeric Results and Plots

Power Comparison

Hypotheses:  $H_0: \mu = \mu_0$  vs.  $H_1: \mu \neq \mu_0$   
 Simulated Distribution: Exponential

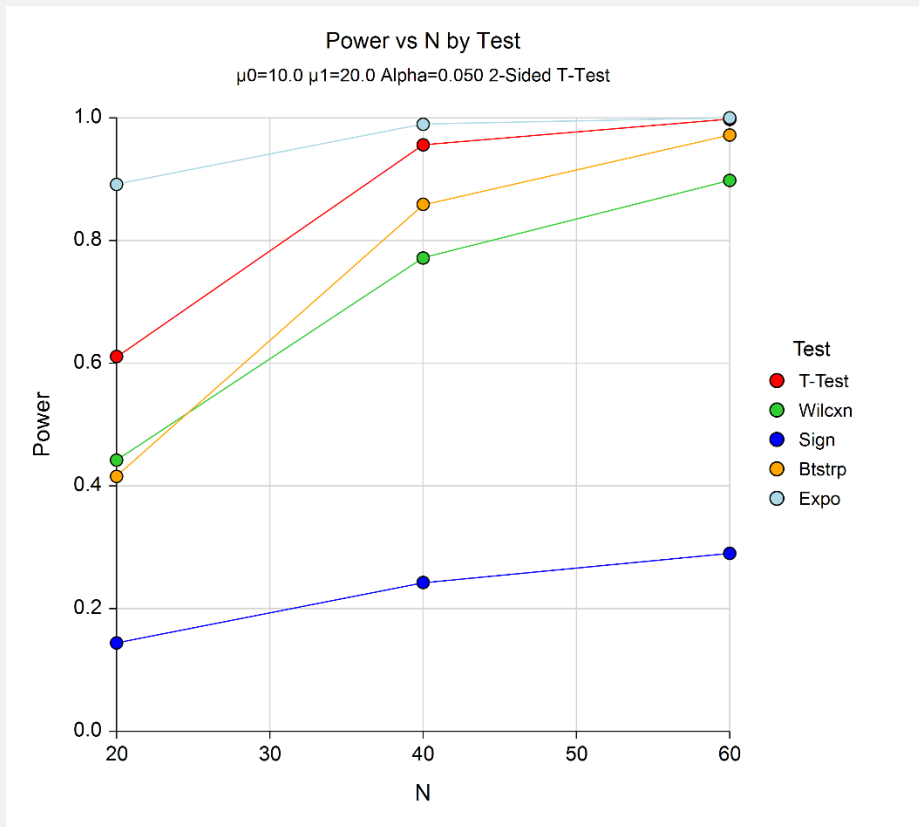
N	$\mu_0$	$\mu_1$	$\sigma$	Target Alpha	T-Test Power	Wilcxn Power	Sign Power	Btstrp Power	Expo Power
20	10.0	20.0	20.0	0.050	0.6110	0.4420	0.1440	0.4155	0.8915
40	10.0	20.0	20.0	0.050	0.9560	0.7715	0.2425	0.8590	0.9895
60	10.0	20.0	20.0	0.050	0.9975	0.8980	0.2900	0.9720	1.0000

Alpha Comparison

Hypotheses:  $H_0: \mu = \mu_0$  vs.  $H_1: \mu \neq \mu_0$   
 Simulated Distribution: Exponential

N	$\mu_0$	$\mu_1$	$\sigma$	Target Alpha	T-Test Alpha	Wilcxn Alpha	Sign Alpha	Btstrp Alpha	Expo Alpha
20	10.0	20.0	20.0	0.050	0.082	0.142	0.203	0.058	0.053
40	10.0	20.0	20.0	0.050	0.073	0.193	0.361	0.059	0.043
60	10.0	20.0	20.0	0.050	0.060	0.270	0.457	0.048	0.046

Chart Section



As would be expected for exponential data, the exponential test performs the best. The bootstrap test performs nearly as well for larger sample sizes. The other tests fail to achieve the target alpha level. Note that these simulation results will vary from run to run because the samples generated are random. The researcher must now decide which test to use based on her level of confidence in the data being truly exponentially distributed and the size of a sample she can afford to take.