Chapter 216

One-Sample T-Test for Non-Inferiority

Introduction

This procedure provides reports for making inference about the non-inferiority of a response compared to a standard value. The question of interest is whether the mean of the response is no worse than a standard reference value (*S*) based on a specified margin (*M*). Another way of saying this is that if the treatment is actually worse than a standard value, it is only worse by a small, acceptable value called the *margin*.

Two different test statistics may be used: the parametric one-sample T-Test and the nonparametric Wilcoxon Signed-Rank Test.

Technical Details

Suppose you want to evaluate the non-inferiority of a continuous random variable, *X*, as compared to a standard value, *S*. Assume that *n* observations, X_k, $k = 1, 2, …, n$ are available.

Non-Inferiority Test

This discussion is based on the book by Rothmann, Wiens, and Chan (2012) which discusses the two-independent sample case. These results are easily applied to the one-sample case.

Assume that **higher values are better**, *μ* represents the population mean of the response, *S* is a standard reference value, and that *M* is the positive *non-inferiority margin*. The null and alternative hypotheses are

$$H0: \mu \leq S - M$$
$$H1: \mu > S - M$$

If, on the other hand, we assume that **higher values are worse**, then null and alternative hypotheses are

$$H0: \mu \geq S + M$$
$$H1: \mu < S + M$$

The one-sample t-test usually tests that the mean is zero. The non-inferiority test compares the mean to a non-zero quantity $S - M$ or $S + M$. This test can also be computed using the One-Sample T-Test procedure.

Assumptions

This section describes the assumptions that are made when you use each of the tests of this procedure. The key assumption relates to normality or non-normality of the data. One of the reasons for the popularity of the t-test is its robustness in the face of assumption violation. Unfortunately, in practice it often happens that more than one assumption is not met. Hence, take the steps to check the assumptions before you make important decisions based on these tests. There are reports in this procedure that permit you to examine the assumptions, both visually and through assumptions tests.

One-Sample T-Test Assumptions

The assumptions of the one-sample t-test are:

1. The data are continuous (not discrete).

2. The data follow the normal probability distribution.

3. The sample is a simple random sample from its population. Each individual in the population has an equal probability of being selected in the sample.

Wilcoxon Signed-Rank Test Assumptions

The assumptions of the Wilcoxon signed-rank test are as follows:

1. The data are continuous (not discrete).

2. The distribution of the data is symmetric.

3. The data are mutually independent.

4. The data all have the same median.

5. The measurement scale is at least interval.

Data Structure

For this procedure, the data are entered as a single column and specified as a response variable.

X1
26
48
19
38
16
16
54
56
87

Procedure Options

This section describes the options available in this procedure.

Variables Tab

This option specifies the variables that will be used in the analysis.

Response Variable

Response Variable

Select (or enter) the name or column number of the response variable to analyze.

Non-Inferiority Test Options

Higher Values Are

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

- **Better**

  If higher values are better the null hypothesis is $\mu \leq$ (Standard Value) - (Non-Inferiority Margin) and the alternative hypothesis is $\mu >$ (Standard Value) - (Non-Inferiority Margin).

- **Worse**

  If higher values are worse the null hypothesis is $\mu \geq$ (Standard Value) + (Non-Inferiority Margin) and the alternative hypothesis is $\mu <$ (Standard Value) + (Non-Inferiority Margin).

Standard Value

This is the known standard or reference value against which the one-sample response will be compared.

Non-Inferiority Margin

Enter the desired value of the non-inferiority margin. The scale of this value is the same as the data values. This value should be positive. (The correct sign will be applied when the null and alternative hypotheses are created based on the selection for "Higher Values Are".)

Reports Tab

The options on this panel specify which reports will be included in the output.

Descriptive Statistics and Confidence Intervals

Descriptive Statistics

This section reports the count, mean, standard deviation, standard error, and mean for the specified variable.

Confidence Level

This value is used to construct the confidence interval of the mean in the descriptive statistics report. Typical confidence levels are 90%, 95%, and 99%, with 95% being the most common.

Tests

Alpha

This is the significance level of the non-inferiority tests. A value of 0.05 is popular. Since this is a one-sided test, the value of 0.025 is often used. Typical values range from 0.001 to 0.200.

Tests – Parametric

T-Test

This report provides the results of the common (parametric) one-sample T-Test. This test should only be used when the data follow the normal distribution.

Tests – Nonparametric

Wilcoxon Signed-Rank Test

This nonparametric test makes use of the sign and the magnitude of the rank of the differences (sample values minus the hypothesized value). It is one of the most commonly used nonparametric alternatives to the one-sample t-test.

There are 3 different tests that can be conducted:

- **Exact Test**

  The exact test can be calculated if there are no ties. This test is recommended when there are no ties.

- **Normal Approximation Test**

  The normal approximation method may be used to approximate the distribution of the sum of ranks when the sample size ≥ 10.

- **Normal Approximation Test with Continuity Correction**

  The normal approximation with continuity correction may be used to approximate the distribution of the sum of ranks when the sample size ≥ 10.

Assumptions

Shapiro-Wilk Test for Normality

This section reports a Shapiro-Wilk normality test.

Assumptions Alpha

This is the significance level of the Shapiro-Wilk normality test. A value of 0.05 is recommended. Typical values range from 0.001 to 0.200.

Report Options Tab

The options on this panel control the label and decimal options of the report.

Report Options

Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

Decimal Places

Means and C.I. Limits – Test Statistics

These options specify the number of decimal places used in the reports. If one of the Auto options is used, the ending zero digits are not shown. For example, if "Significant Digits (Up to 7)" is chosen,

0.0500 is displayed as 0.05 and 1.314583689 is displayed as 1.314584.

The output formatting system is not designed to accommodate (Up to 13), and if chosen, this will likely lead to lines that run on to a second line. This option is included, however, for the rare case when a very large number of decimals is needed.

Plots Tab

The options on this panel control the inclusion and appearance of the plots.

Select Plots

Histogram … Average-Difference Plot

Check the boxes to display the plot. Click the plot format button to change the plot settings.

Example 1 – Non-Inferiority Test for a Single Sample

This section presents an example of how to test non-inferiority for a single sample. Suppose a new treatment method is being investigated and a measurement was made on each of 100 subjects using the new treatment (Method1). The data are in the **Bland-Altman** dataset. The researchers wish to determine if the treatment response is no more than 0.5 less than the known standard value of 51 on average.

You may follow along here by making the appropriate entries or load the completed template **Example 1** by clicking on Open Example Template from the File menu of the One-Sample T-Test for Non-Inferiority window.

1   Open the Bland-Altman dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Bland-Altman.NCSS**.
- Click **Open**.

2   Open the One-Sample T-Test for Non-Inferiority window.

- Using the Analysis menu or the Procedure Navigator, find and select the **One-Sample T-Test for Non-Inferiority** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3   Specify the variables.

- Select the **Variables tab**. (This is the default.)
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Method1** from the list of variables and then click **Ok**. "Method1" will appear in this box.
- Leave **Higher Values Are** at **Better**.
- Set **Standard Value** to **51**.
- Change the **Non-Inferiority Margin** to **0.5**.

4   Specify the reports.

- Select the **Reports tab**.
- Leave all checked reports as checked.
- Check the **Wilcoxon Signed-Rank Test** check box. Leave all sub-boxes checked.

5   Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

The following reports and charts will be displayed in the Output window.

Descriptive Statistics

Descriptive Statistics							
Variable	Count	Mean	Standard Deviation of Data	Standard Error of Mean	T*	95.0% LCL of Mean	95.0% UCL of Mean
Method1	100	50.72	28.30893	2.830893	1.9842	45.10289	56.3371

This report provides basic descriptive statistics and confidence interval for the response variable.

Variable

The name of the variable for which the descriptive statistics are listed here.

Count

The count gives the number of non-missing values. This value is often referred to as the group sample size or *n*.

Mean

This is the average of the data values.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Standard Deviation

The sample standard deviation is the square root of the variance. It is a measure of dispersion based on squared distances from the mean for the variables listed.

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}}$$

Standard Error

This is the estimated standard deviation of the distribution of sample means for an infinite population.

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

T*

This is the t-value used to construct the confidence limits. It is based on the degrees of freedom and the confidence level.

LCL and UCL of the Mean

This is the lower limit of an interval estimate of the mean based on a Student's t distribution with n - 1 degrees of freedom. This interval estimate assumes that the population standard deviation is not known and that the data are normally distributed. The confidence interval formula is

$$\bar{x} \pm T_{df} \cdot SE_{\bar{x}}$$

One-Sample T-Test for Non-Inferiority

One-Sample T-Test for Non-Inferiority
Higher Values are Better
Non-Inferiority Hypothesis: (Method1) > 51 - 0.5

Alternative Hypothesis	Mean	Standard Error	T-Statistic	DF	Prob Level	Conclude Non-Inferiority at α = 0.050?
μ > 50.5	50.72	2.830893	0.0777	99	0.46911	No

This report shows the parametric one-sample T-Test for non-inferiority. Since the Prob Level is greater than the designated value of alpha (0.05), the null hypothesis of inferiority cannot be rejected in favor of the alternative hypothesis of non-inferiority. This test assumes that the data are normal.

Alternative Hypothesis

Assume that μ represents the mean of the response data values, S is a standard reference value, and that M is the positive *non-inferiority margin*. If **higher values are better**, then the null and alternative hypotheses are

$$H0: \mu \leq S - M$$
$$H1: \mu > S - M$$

If **higher values are worse**, then null and alternative hypotheses are

$$H0: \mu \geq S + M$$
$$H1: \mu < S + M$$

In practice, the alternative hypothesis should be chosen in advance.

Mean

This is the average of the data values.

Standard Error

This is the estimated standard deviation of the distribution of sample means for an infinite population.

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

T-Statistic

The T-Statistic is the value used to produce the p-value (Prob Level) based on the T distribution. The formula for the T-Statistic is:

$$T = \frac{\bar{x} - (S - M)}{SE_{\bar{x}}}$$

DF

The degrees of freedom define the T *distribution* upon which the probability values are based. The formula for the degrees of freedom is:

$$df = n - 1$$

Prob Level

The probability level, also known as the p-value or significance level, is the probability that the test statistic will take a value at least as extreme as the observed value, assuming that the null hypothesis is true. If the p-value is less than the prescribed α, in this case 0.05, the null hypothesis is rejected in favor of the alternative hypothesis. Otherwise, there is not sufficient evidence to reject the null hypothesis.

Conclude Non-Inferiority at α = 0.050?

This column indicates whether or not the null hypothesis is rejected, in favor of the alternative hypothesis, based on the p-value and chosen α. A test in which the null hypothesis is rejected is sometimes called *significant*.

Wilcoxon Signed-Rank Test for Non-Inferiority

Wilcoxon Signed-Rank Test for Non-Inferiority
Higher Values are Better
Non-Inferiority Hypothesis: (Median of Method1) > 51 - 0.5

Sum of Ranks (W)	Mean of W	Std Dev of W	Number of Zeros	Number Sets of Ties	Multiplicity Factor
2551	2525	290.8101	0	31	816

Test Type	Alternative Hypothesis	Z-Value	Prob Level	Conclude Non-Inferiority at α = 0.050?
Exact*	Median > 50.5			
Normal Approximation	Median > 50.5	0.0894	0.46438	No
Normal Approx. with C.C.	Median > 50.5	0.0877	0.46506	No

*The Exact Test is provided only when there are no ties.

This report shows the nonparametric Wilcoxon Signed-Rank Test for non-inferiority. Since the Prob Level is greater than the designated value of alpha (0.05), the null hypothesis of inferiority cannot be rejected in favor of the alternative hypothesis of non-inferiority. This test required no assumption that the data are normal.

Sum of Ranks (W)

The basic statistic for this test is the sum of the positive ranks, ΣR_+ (The sum of the positive ranks is chosen arbitrarily. The sum of the negative ranks could equally be used). This statistic is called W.

$$W = \sum R_+$$

Mean of W

This is the mean of the sampling distribution of the sum of ranks for a sample of n items.

$$\mu_W = \frac{n(n+1) - d_0(d_0 + 1)}{4}$$

where d_0 is the number of zero differences.

Std Dev of W

This is the standard deviation of the sampling distribution of the sum of ranks. Here t_i represents the number of times the ith value occurs.

$$s_W = \sqrt{\frac{n(n+1)(2n+1) - d_0(d_0+1)(2d_0+1)}{24} - \frac{\sum t_i^3 - \sum t_i}{48}}$$

where d_0 is the number zero differences, t_i is the number of absolute differences that are tied for a given non-zero rank, and the sum is over all sets of tied ranks.

Number of Zeros

This is the number of times that the difference between the observed data and the hypothesized value is zero. The zeros are used in computing ranks, but are not considered positive ranks or negative ranks.

Number Sets of Ties

The treatment of ties is to assign an average rank for the particular set of ties. This is the number of sets of ties that occur in the data, including ties at zero.

Multiplicity Factor

This is the correction factor that appeared in the standard deviation of the sum of ranks when there were ties.

Test Type

This is the type of test that is being reported on the current row. The Exact Test is provided only when there are no ties.

Alternative Hypothesis

For the Wilcoxon signed-rank test, the null and alternative hypotheses relate to the median. The left-tail alternative is represented by Median < S + M (i.e., H_a: Median < S + M) while the right-tail alternative is depicted by Median > S - M.

Exact Probability: Prob Level

This is an exact p-value for this statistical test, assuming no ties. The p-value is the probability that the test statistic will take on a value at least as extreme as the actually observed value, assuming that the null hypothesis is true. If the p-value is less than α, say 5%, the null hypothesis is rejected. If the p-value is greater than α, the null hypothesis is accepted.

Exact Probability: Reject H0 (α = 0.050)

This is the conclusion reached about the null hypothesis. It will be to either fail to reject H_0 or reject H_0 at the assigned level of significance.

Approximations with (and without) Continuity Correction: Z-Value

Given the sample size is at least ten, a normal approximation method may be used to approximate the distribution of the sum of ranks. Although this method does correct for ties, it does not have the continuity correction factor. The *z* value is as follows:

$$z = \frac{W - \mu_W}{\sigma_W}$$

If the correction factor for continuity is used, the formula becomes:

$$z = \frac{W - \mu_W \pm \frac{1}{2}}{\sigma_W}$$

Approximations with (and without) Continuity Correction: Prob Level

This is the *p*-value for the normal approximation approach for the Wilcoxon signed-rank test. The *p*-value is the probability that the test statistic will take a value at least as extreme as the actually observed value, assuming that the null hypothesis is true. If the *p*-value is less than *α*, say 5%, the null hypothesis is rejected. If the *p*-value is greater than *α*, the null hypothesis is accepted.

Approximations with (and without) Continuity Correction: Reject H0 (α = 0.050)

This is the conclusion reached about the whether to reject null hypothesis. It will be either Yes or No at the given level of significance.
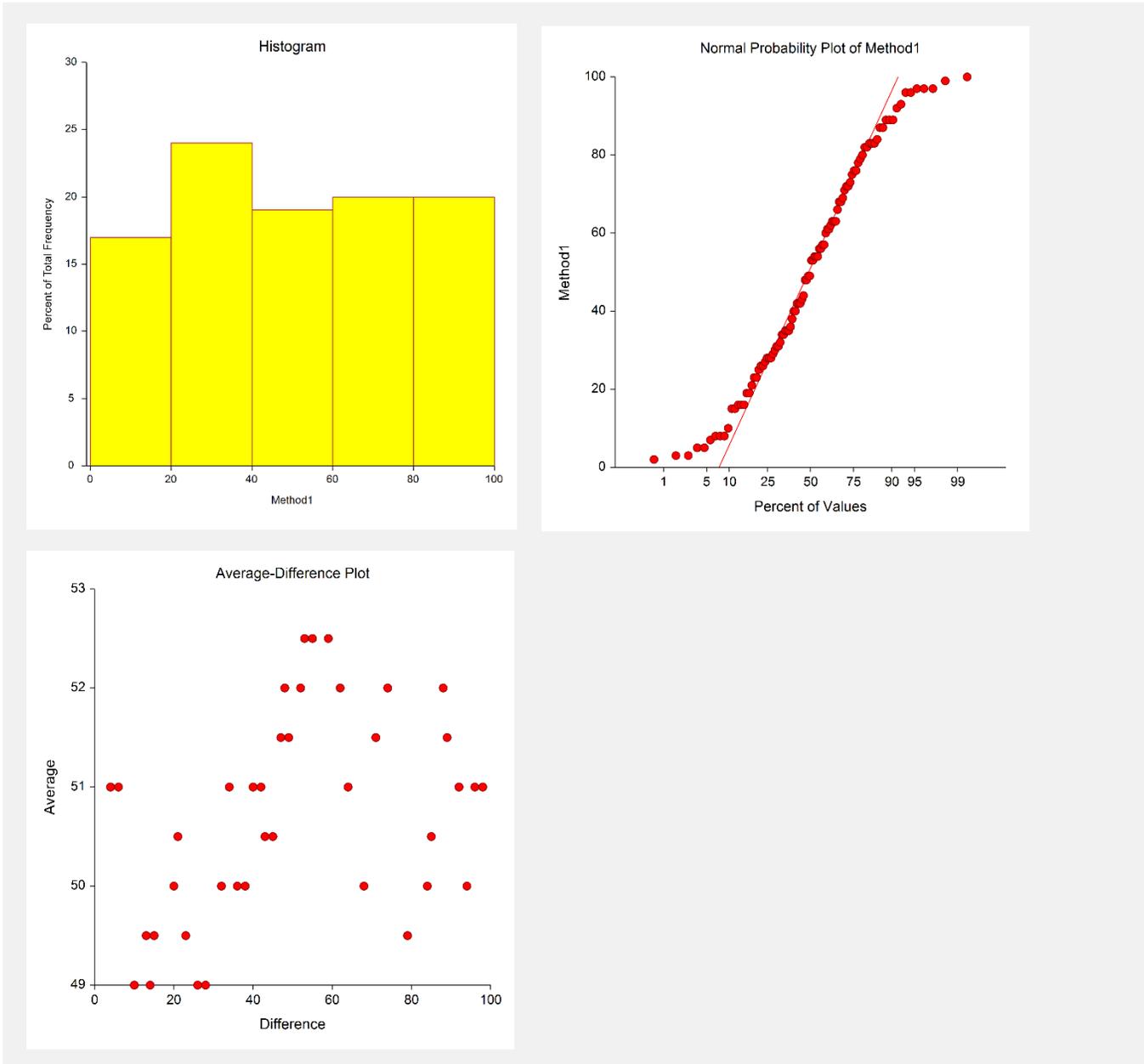
Test of Assumptions

Test of the Data Normality Assumption

Assumption	Value	Prob Level	Decision (α = 0.050)
Shapiro-Wilk	0.9580	0.00292	Reject normality

The main assumption when using the t-test is that the data come from a normal distribution. The normality assumption can be checked statistically by the Shapiro-Wilk normality test and visually by the histogram or normal probability plot.

This section reports the results of a diagnostic test to determine if the data are normal. In this case, they are not, probably because of the outliers that were present. This would indicate that the Wilcoxon Signed-Rank Test would be the better test to use.

Plots Section

Histogram

The nonparametric tests need the assumption of symmetry, and these two graphic tools can provide that information. Since the histogram's shape is impacted by the number of classes or bins and the width of the bins, the best choice is to trust the density trace, which is a smoothed histogram. If the distribution of data is symmetrical but not normal, proceed with the nonparametric test.

Normal Probability Plot

If any of the observations fall outside the confidence bands (if shown), the data are not normal. The goodness-of-fit test mentioned earlier should confirm this fact statistically. If only one observation falls outside the confidence bands and the remaining observations hug the straight line, there may be an outlier. If the data were normal, we would see the points falling along a straight line. Note that the confidence bands are based on large-sample formulas. They may not be accurate for small samples.

Average-Difference Plot

This average-difference plot is designed to detect a lack of symmetry in the data. This plot is constructed from the original data. Let $D(i)$ represent the ith ordered difference. Pairs of these sorted differences are considered, with the pairing being done as you move toward the middle from either end. That is, consider the pairs $D(1)$ and $D(n)$, $D(2)$ and $D(n-1)$, $D(3)$ and $D(n-2)$, etc. Plot the average versus the difference of each of these pairs. Your plot will have about $n/2$ points, depending on whether n is odd or even. If the data are symmetric, the average of each pair will be the median and the difference between each pair will be zero.

Symmetry is an important assumption for the t-test. A perfectly symmetric set of data should show a vertical line of points hitting the horizontal axis at the value of the median. Departures from symmetry would deviate from this standard.