Chapter 198

Two-Sample T-Test for Non-Inferiority

Introduction

This procedure provides reports for making inference about the non-inferiority of a treatment mean compared to a control mean from data taken from independent groups. The question of interest is whether the treatment mean is better than or, at least, no worse than the control mean. Another way of saying this is that if the treatment mean is actually worse than the control mean, it is only worse by a small, acceptable value called the *margin*.

Three different test statistics may be used: two-sample t-test, the Aspin-Welch unequal-variance t-test, and the nonparametric Mann-Whitney U (or Wilcoxon Rank-Sum) test.

Technical Details

Suppose you want to evaluate the non-inferiority of a continuous random variable X_T as compared to a second random variable X_C using data on each variable taken on the different subjects. Assume that n_T observations (X_{Tk}), $k = 1, 2, \ldots, n_T$ are available from the treatment group and that n_C observations (X_{Ck}), $k = 1, 2, \ldots, n_C$ are available from the control group.

Non-Inferiority Test

This discussion is based on the book by Rothmann, Wiens, and Chan (2012) which discusses the two-independent sample case. Assume that higher values are better, that μ_T and μ_C represent the means of the two variables, and that M is the positive *non-inferiority margin.* The null and alternative hypotheses when the **higher values are better** are

$$\text{H0: } (\mu_T - \mu_C) \le -M$$
$$\text{H1: } (\mu_T - \mu_C) > -M$$

or

$$\text{H0: } \mu_T \le \mu_C - M$$
$$\text{H1: } \mu_T > \mu_C - M$$

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

$$\text{H0: } (\mu_T - \mu_C) \ge M$$
$$\text{H1: } (\mu_T - \mu_C) < M$$

or

$$\text{H0: } \mu_T \ge \mu_C + M$$
$$\text{H1: } \mu_T < \mu_C + M$$

The two-sample t-test is usually employed to test that the mean difference is zero. The non-inferiority test is a one-sided two-sample t-test that compares the difference to a non-zero quantity, *M*. One-sided editions of the Aspin-Welch unequal-variance t-test, and the Mann-Whitney U (or Wilcoxon Rank-Sum) nonparametric test are also optionally available.

Data Structure

The data may be entered in two formats, as shown in the two examples below. The examples give the yield of corn for two types of fertilizer. The first format, shown in the first table, is the case in which the responses for each group are entered in separate columns. That is, each variable contains all responses for a single group. In the second format the data are arranged so that all responses are entered in a single column. A second column, referred to as the *grouping variable*, contains an index that gives the group (A or B) to which the row of data belongs.

In most cases, the second format is more flexible. Unless there is some special reason to use the first format, we recommend that you use the second.

Two Response Variables

Yield A	Yield B
452	546
874	547
554	774
447	465
356	459
754	665
558	467
574	365
664	589
682	534
547	456
435	651
245	654
	665
	546
	537

Grouping and Response Variables

Fertilizer	Yield
B	546
B	547
B	774
B	465
B	459
B	665
B	456
.	.
.	.
A	452
A	874
A	554
A	447
A	356
A	754
A	558
A	574
A	664
.	.
.	.

Procedure Options

This section describes the options available in this procedure.

Variables Tab

These options specify the variables that will be used in the analysis as well as the non-inferiority margin.

Variables

Data Input Type

In this procedure, there are two ways to organize the data. Select the type that reflects the way your data are presented on the spreadsheet.

- **Response Variable and Group Variable**

  In this scenario, the response data is in one column and the groups are defined in another column of the same length. For example, you might have

Response	Group
12	1
35	1
19	1
24	1
26	2
44	2
36	2
33	2

  If the group variable has more than two levels, a comparison is made among each pair of levels.

- **Two Variables with Response Data in each Variable**

  In this selection, the data for each group are in separate columns. You are given two boxes to select the treatment group variable and the control group variable.

Treatment	Control
12	26
35	44
19	36
24	33

Variables – Data Input Type: Response Variable and Group Variable

For this input type, the group data are in one column and the response data are in another column of the same length.

Response	Group
12	1
35	1
19	1
24	1
26	2
44	2
36	2
33	2

Response Variable

Specify the variable containing the response data.

Group Variable

Specify the variable defining the grouping of the response data. If the group variable has more than two levels, a comparison is made among each pair of levels.

Variables – Data Input Type: Two Variables with Response Data in each Variable

For this data input type, the data for each group are in separate columns. The number of values in each column need not be the same.

Treatment	Control
12	26
35	44
19	36
24	33
27	
32	

Treatment Variable

Specify the variable that contains the treatment data.

Control Variable

Specify the variable that contains the control data.

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 H0: Treatment Mean ≤ Control Mean - Margin and the alternative hypothesis is H1: Treatment Mean > Control Mean - Margin. That is, the treatment mean is no more than a small margin below the control mean.

- **Worse**

  If higher values are worse the null hypothesis is H0: Treatment Mean ≥ Control Mean + Margin and the alternative hypothesis is H1: Treatment Mean < Control Mean + Margin. That is, the treatment mean is no more than a small margin above the control mean.

Non-Inferiority Margin

Enter the desired value of the non-inferiority margin. The scale of this value is the same as the data values. For example, if the control mean is historically equal to 67, a realistic margin might be 5% or 3.35.

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” above.).

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 means, medians, standard deviations, standard errors, and confidence intervals of each variable and the mean difference.

Confidence Level

This confidence level is used for the descriptive statistics confidence intervals of each group, as well as for the confidence interval of the mean difference. 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 test. 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

Equal-Variance T-Test

This provides the results of the non-inferiority test under the assumption that the two group variances are equal.

Unequal-Variance T-Test

This provides the results of the non-inferiority test under the assumption that the two group variances are not equal.

Tests – Nonparametric

Mann-Whitney U Test (Wilcoxon Rank-Sum Test)

This test is a nonparametric alternative to the equal-variance t-test for use when the assumption of normality is not valid. This test uses the ranks of the values rather than the values themselves.

There are 3 different tests that can be conducted:

- **Exact Test**

  The exact test can be calculated if there are no ties and the sample size is ≤ 20 in both groups. This test is recommended when these conditions are met.

- **Normal Approximation Test**

  The normal approximation method may be used to approximate the distribution of the sum of ranks when the sample size is reasonably large.

- **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 is reasonably large.

Assumptions

Tests of Assumptions

This section reports normality tests and equal-variance tests.

Assumptions Alpha

This is the significance level of the various tests of normality and equal variance. 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.

Value Labels

If a grouping variable is used, this option lets you indicate how it is labelled.

Decimal Places

Means, Differences, 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 *Auto (Up to 7)* is chosen,

0.0534 is displayed as 0.0534

and

1.314583689 is displayed as 1.314584.

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

Plots Tab

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

Select Plots

Histograms, Probability Plots, and Box Plot

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

Example 1 – Non-Inferiority Test for Two Independent Samples

This section presents an example of how to test non-inferiority. Suppose the current (control) fertilizer has an undesirable impact on the ground water so a replacement (treatment) fertilizer has been developed that does not have this negative impact. The researchers of the new fertilizer want to show that the new fertilizer is not less than a small margin below the current fertilizer. Further suppose that the average corn yield of the current fertilizer is about 550. The researchers want to show that the yield of the new fertilizer is not less than 20% below the current type. That is, the non-inferiority margin is 20% of 550 which is 110.

The data are in the **Corn Yield** dataset. 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 Two-Sample T-Test for Non-Inferiority window.

1 Open the Corn Yield dataset.

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

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

- Using the Analysis menu or the Procedure Navigator, find and select the **Two-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**.
- Set the **Data Input Type** box to **Two Variables with Response Data in each Variable**.
- Double-click in the **Treatment Variable** text box. This will bring up the variable selection window.
- Select **YldA** from the list of variables and then click **Ok**. “YldA” will appear in this box.
- Double-click in the **Control Variable** text box. This will bring up the variable selection window.
- Select **YldB** from the list of variables and then click **Ok**. “YldB” will appear in this box.
- Set the **Higher Values Are** box to **Better**.
- Change the **Non-Inferiority Margin** to **110**.

4 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

Variable	Count	Mean	Standard Deviation of Data	Standard Error of Mean	T*	95% LCL of Mean	95% UCL of Mean
YldA	13	549.3846	168.7629	46.80641	2.1788	447.4022	651.367
YldB	16	557.5	104.6219	26.15546	2.1314	501.7509	613.249

This report provides basic descriptive statistics and confidence intervals for the two variables.

Variable

These are the names of the variables or groups.

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 for each group.

Standard Deviation of Data

The sample standard deviation is the square root of the sample variance. It is a measure of spread.

Standard Error of Mean

This is the estimated standard deviation for the distribution of sample means for an infinite population. It is the sample standard deviation divided by the square root of sample size.

T*

This is the t-value used to construct the confidence interval. If you were constructing the interval manually, you would obtain this value from a table of the Student's t distribution with *n* - 1 degrees of freedom.

LCL 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.

UCL of the Mean

This is the upper limit of the interval estimate for the mean based on a t distribution with *n* - 1 degrees of freedom.

Confidence Intervals for the Mean Difference

Variance Assumption	DF	Mean Difference	Standard Deviation	Standard Error	T*	95% LCL of Difference	95% UCL of Difference
Equal	27	-8.115385	136.891	51.11428	2.0518	-112.9932	96.76247
Unequal	19.17	-8.115385	198.5615	53.61855	2.0918	-120.2734	104.0426

Given that the assumptions of independent samples and normality are valid, this section provides an interval estimate (confidence limits) of the difference between the two means. Results are given for both the equal and unequal variance cases.

DF

The degrees of freedom are used to determine the T distribution from which T* is generated.

For the equal variance case:

$$df = n_T + n_C - 2$$

For the unequal variance case:

$$df = \frac{\left(\frac{s_T^2}{n_T} + \frac{s_C^2}{n_C}\right)^2}{\frac{\left(\frac{s_T^2}{n_T}\right)^2}{n_T - 1} + \frac{\left(\frac{s_C^2}{n_C}\right)^2}{n_C - 1}}$$

Mean Difference

This is the difference between the sample means, $\overline{X}_T - \overline{X}_C$.

Standard Deviation

In the equal variance case, this quantity is:

$$s_{\overline{X}_T-\overline{X}_C} = \sqrt{\frac{(n_T-1)s_T^2+(n_C-1)s_C^2}{n_T-n_C-2}}$$

In the unequal variance case, this quantity is:

$$s_{\overline{X}_T-\overline{X}_C} = \sqrt{s_T^2+s_C^2}$$

Standard Error

This is the estimated standard deviation of the distribution of differences between independent sample means.

For the equal variance case:

$$SE_{\overline{X}_T-\overline{X}_C} = \sqrt{\left(\frac{(n_T-1)s_T^2+(n_C-1)s_C^2}{n_T-n_C-2}\right)\left(\frac{1}{n_T}+\frac{1}{n_C}\right)}$$

For the unequal variance case:

$$SE_{\overline{X}_T-\overline{X}_C} = \sqrt{\frac{s_T^2}{n_T}+\frac{s_C^2}{n_C}}$$

T*

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

Lower and Upper Confidence Limits

These are the confidence limits of the confidence interval for $\mu_T - \mu_C$. The confidence interval formula is

$$\overline{X}_T - \overline{X}_C \pm T_{df}^* \cdot SE_{\overline{X}_T-\overline{X}_C}$$

The equal-variance and unequal-variance assumption formulas differ by the values of T* and the standard error.

Descriptive Statistics for the Median

Variable	Count	Median	95% LCL of Median	95& UCL of Median
YldA	13	554	435	682
YldB	16	546	465	651

This report provides the medians and corresponding confidence intervals for the medians of each group.

Variable

These are the names of the variables or groups.

Count

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

Median

The median is the 50th percentile of the group data, using the AveXp(n+1) method. The details of this method are described in the Descriptive Statistics chapter under Percentile Type.

LCL and UCL

These are the lower and upper confidence limits of the median. These limits are exact and make no distributional assumptions other than a continuous distribution. No limits are reported if the algorithm for this interval is not able to find a solution. This may occur if the number of unique values is small.

Equal-Variance T-Test for Non-Inferiority

Equal-Variance T-Test for Non-Inferiority
Higher Values are Better
Non-Inferiority Hypothesis: (YldA) > (YldB) - 110

Alternative Hypothesis	Mean Difference	Standard Error	T-Statistic	DF	Prob Level	Conclude Non-Inferiority at α = 0.05?
μT > μC - 110	-8.115385	51.11428	1.9933	27	0.02821	Yes

This report shows the non-inferiority test for the equal-variance assumption. Since the Prob Level is less than the designated value of alpha (0.05), the null hypothesis of inferiority is rejected and the alternative hypothesis of non-inferiority is concluded.

Aspin-Welch Unequal-Variance T-Test for Non-Inferiority

Aspin-Welch Unequal-Variance T-Test for Non-Inferiority
Higher Values are Better
Non-Inferiority Hypothesis: (YldA) > (YldB) - 110

Alternative Hypothesis	Mean Difference	Standard Error	T-Statistic	DF	Prob Level	Conclude Non-Inferiority at α = 0.05?
μT > μC - 110	-8.115385	53.61855	1.9002	19.17	0.03628	Yes

This report shows the non-inferiority test for the unequal-variance assumption. Since the Prob Level is again less than the designated value of alpha (0.05), the null hypothesis of inferiority is rejected and the alternative hypothesis of non-inferiority is concluded.

Mann-Whitney U or Wilcoxon Rank-Sum Location Difference Test for Non-Inferiority

Mann-Whitney U or Wilcoxon Rank-Sum Location Difference Test for Non-Inferiority
Higher Values are Better
Non-Inferiority Hypothesis: (YldA) > (YldB) - 110

Variable	Mann-Whitney U	Sum of Ranks (W)	Mean of W	Std Dev of W
YldA	150.5	241.5	195	22.79508
YldB	57.5	193.5	240	22.79508

Number of Sets of Ties = 3, Multiplicity Factor = 18

Test Type	Alternative Hypothesis†	Z-Value	Prob Level	Conclude Non-Inferiority at α = 0.050?
Exact*	LocT > LocC - 110			
Normal Approximation	LocT > LocC - 110	2.0399	0.02068	Yes
Normal Approx. with C.C.	LocT > LocC - 110	2.0180	0.02180	Yes

† "LocT" and "LocC" refer to the location parameters of the treatment and control distributions, respectively.
* The Exact Test is provided only when there are no ties and the sample size is ≤ 20 in both groups.

This report shows the non-inferiority test based on the Mann-Whitney U statistic. This test is documented in the Two-Sample T-Test chapter.

Tests of Assumptions

Tests of the Normality Assumption for YldA

Test Name	Test Statistic	Reject H0 of Prob Level	Normality Decision (α = 0.05)
Shapiro-Wilk	0.9843	0.99420	No
Skewness	0.2691	0.78785	No
Kurtosis	0.3081	0.75803	No
Omnibus (Skewness or Kurtosis)	0.1673	0.91974	No

Tests of the Normality Assumption for YldB

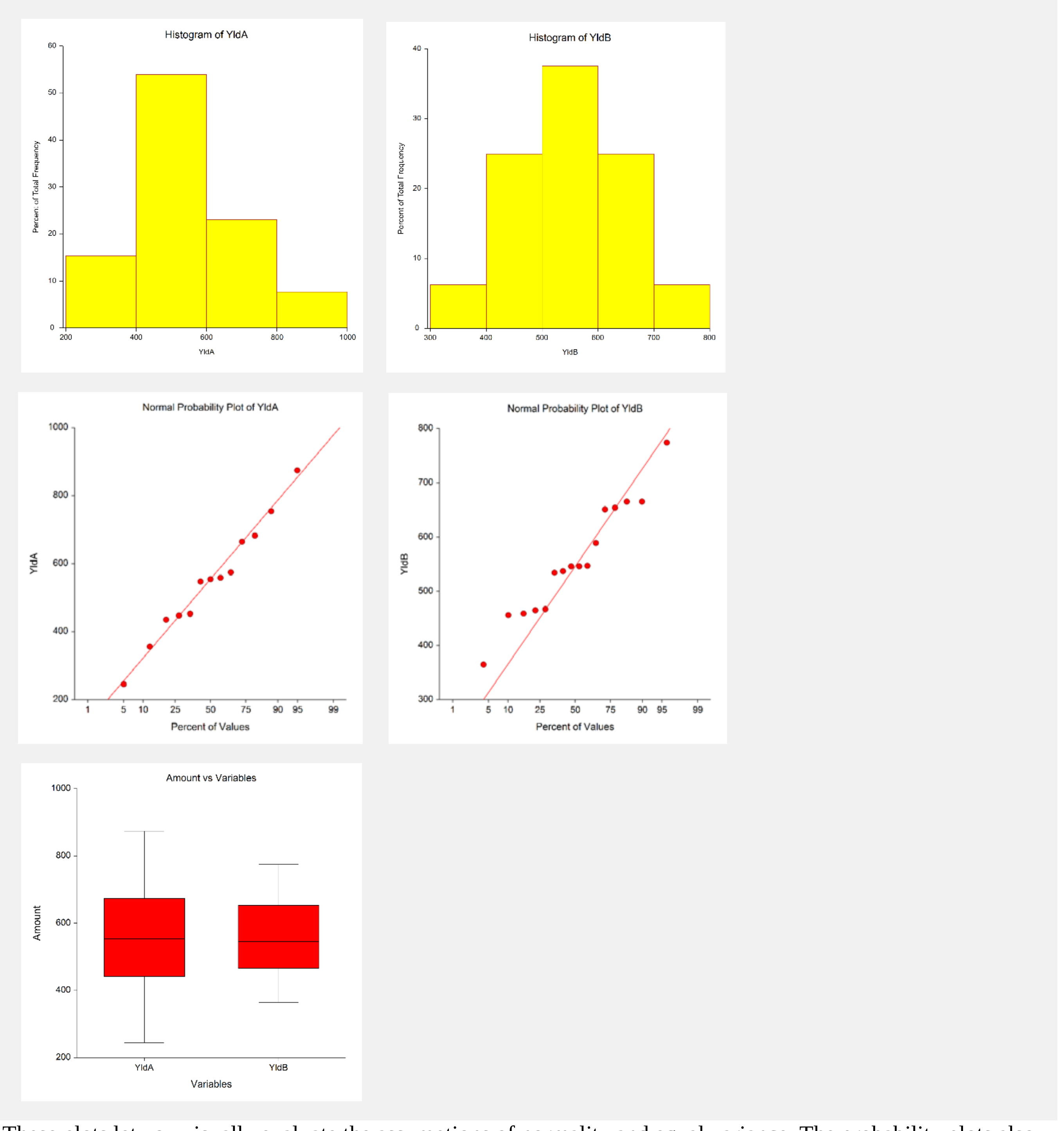
Test Name	Test Statistic	Reject H0 of Prob Level	Normality Decision (α = 0.05)
Shapiro-Wilk	0.9593	0.64856	No
Skewness	0.4587	0.64644	No
Kurtosis	0.1291	0.89726	No
Omnibus (Skewness or Kurtosis)	0.2271	0.89267	No

Tests of the Equal Variance Assumption

Test Name	Test Statistic	Reject H0 of Prob Level	Equal Variances Decision (α = 0.05)
Variance-Ratio	2.6020	0.08315	No
Modified-Levene	1.9940	0.16935	No

This section reports the results of diagnostic tests to determine if the data are normal and the variances are close to being equal. The details of these tests are given in the Descriptive Statistics chapter.

Evaluation of Assumptions Plots



These plots let you visually evaluate the assumptions of normality and equal variance. The probability plots also let you see if outliers are present in the data.