Chapter 415

Multivariate Analysis of Variance (MANOVA)

Introduction

Multivariate analysis of variance (MANOVA) is an extension of common analysis of variance (ANOVA). In ANOVA, differences among various group means on a single-response variable are studied. In MANOVA, the number of response variables is increased to two or more. The hypothesis concerns a comparison of vectors of group means. When only two groups are being compared, the results are identical to Hotelling's T^2 procedure.

The multivariate extension of the F-test is not completely direct. Instead, several test statistics are available, such as Wilks' Lambda and Lawley's trace. The actual distributions of these statistics are difficult to calculate, so we rely on approximations based on the F-distribution.

Technical Details

A MANOVA has one or more factors (each with two or more levels) and two or more dependent variables. The calculations are extensions of the general linear model approach used for ANOVA.

Unlike the univariate situation in which there is only one statistical test available (the F-ratio), the multivariate situation provides several alternative statistical tests. We will describe these tests in terms of two matrices, H and E. H is called the *hypothesis matrix* and E is the *error matrix*. These matrices may be computed using a number of methods. In **NCSS**, we use the standard general linear models (GLM) approach in which a sum of squares and cross-products matrix is computed. This matrix is based on the dependent variables and independent variables generated for each degree of freedom in the model. It may be partitioned according to the terms in the model.

MANOVA Test Statistics

For a particular p-variable multivariate test, assume that the matrices *H* and *E* have *h* and *e* degrees of freedom, respectively. Four tests may be defined as follows. See Seber (1984) for details. Let θ_i, ϕ_i, and λ_i be the eigenvalues of $H(E+H)^{-1}$, HE^{-1}, and $E(E+H)^{-1}$ respectively. Note that these eigenvalues are related as follows:

$$\theta_i = 1 - \lambda_i = \frac{\phi_i}{1+\phi_i}$$

$$\phi_i = \frac{\theta_i}{1-\theta_i} = \frac{1-\lambda_i}{\lambda_i}$$

$$\lambda_i = 1 - \theta_i = \frac{1}{1+\phi_i}$$

Wilks' Lambda

Define Wilks' Lambda as follows:

$$\Lambda_{p,h,e} = \frac{|E|}{|E+H|}$$

$$= \prod_{j=1}^{p} \left(1-\theta_j\right)$$

with $e \geq p$.

The following approximation based on the F-distribution is used to determine significance levels:

$$F_{ph,ft-g} = \frac{(ft-g)\left(1-\Lambda^{1/t}\right)}{ph\Lambda^{1/t}}$$

where

$$f = e - \frac{1}{2}(p-h+1)$$

$$g = \frac{ph-2}{2}$$

$$t = \begin{cases} \sqrt{\dfrac{p^2h^2-4}{p^2+h^2-5}} & \text{if } p^2+h^2-5>0 \\ 1 & \text{otherwise} \end{cases}$$

This approximation is exact if *p* or *h* ≥ 2.

Lawley - Hotelling Trace

The trace statistic, T_g^2, is defined as follows:

$$T_g^2 = e \sum_{j=1}^{s} \phi_j$$

where

$$s = \min(p, h)$$

The following approximation based on the F-distribution is used to determine significance levels:

$$F_{a,b} = \frac{T_g^2}{ce}$$

where

$$a = ph$$

$$b = 4 + (a + 2)/(B - 1)$$

$$c = \frac{a(b - 2)}{b(e - p - 1)}$$

$$B = \frac{(e + h - p - 1)(e - 1)}{(e - p - 3)(e - p)}$$

Pillai's Trace

Pillai's trace statistic, $V^{(s)}$, is defined as follows:

$$V^{(s)} = \sum_{j=1}^{s} \theta_j = \operatorname{tr}(H(E + H)^{-1})$$

where

$$s = \min(p, h)$$

The following approximation based on the F-distribution is used to determine significance levels:

$$F_{s(2m+s+1),s(2n+s+1)} = \frac{(2n + s + 1)V^{(s)}}{(2m + s + 1)(s - V^{(s)})}$$

where

$$s = \min(p, h)$$

$$m = (|p - h| - 1)/2$$

$$n = (e - p - 1)/2$$

Roy's Largest Root

Roy's largest root, ϕ_{max}, is defined as the largest of the ϕ_i's. The following approximation based on the F-distribution is used to determine significance levels:

$$F_{(2_{v1}+2),(2_{v2}+2)} = \frac{2_{v2} + 2}{2_{v1} + 2} \phi_{max}$$

where

$$s = \min(p, h)$$

$$v_1 = (|p - h| - 1)/2$$

$$v_2 = (e - p - 1)/2$$

Which Test to Use

When the hypothesis degrees of freedom, *h*, is one, all four test statistics will lead to identical results. When *h*>1, the four statistics will usually lead to the same result. When they do not, the following guidelines from Tabachnick (1989) may be of some help.

Wilks' Lambda, Lawley's trace, and Roy's largest root are often more powerful than Pillai's trace if *h*>1 and one dimension accounts for most of the separation among groups. Pillai's trace is more robust to departures from assumptions than the other three.

Tabachnick (1989) provides the following checklist for conducting a MANOVA. We suggest that you consider these issues and guidelines carefully.

Assumptions and Limitations

The following assumptions are made when using a MANOVA.

1. The response variables are continuous.
2. The residuals follow the multivariate-normal probability distribution with means equal to zero.
3. The variance-covariance matrices of each group of residuals are equal.
4. The individuals are independent.

Multivariate Normality and Outliers

MANOVA is robust to modest amount of skewness in the data. A sample size that produces 20 degrees of freedom in the univariate F-test is adequate to ensure robustness. Non-normality caused by the presence of outliers can cause severe problems that even the robustness of the test will not overcome. You should screen your data for outliers and run it through various univariate and multivariate normality tests and plots to determine if the normality assumption is reasonable.

Homogeneity of Covariance Matrices

MANOVA makes the assumption that the within-cell (group) covariance matrices are equal. If the design is balanced so that there is an equal number of observations in each cell, the robustness of the MANOVA tests is guaranteed. If the design is unbalanced, you should test the equality of covariance matrices using Box's M test. If this test is significant at less than 0.001, there may be severe distortion in the alpha levels of the tests. You should only use Pillai's trace criterion in this situation.

Linearity

MANOVA assumes linear relationships among the dependent variables within a particular cell. You should study scatter plots of each pair of dependent variables using a different color for each level of a factor. Look carefully for curvilinear patterns and for outliers. The occurrence of curvilinear relationships will reduce the power of the MANOVA tests.

Multicollinearity and Singularity

Multicollinearity occurs when one dependent variable is almost a weighted average of the others. This collinearity may only show up when the data are considered one cell at a time. The R^2-Other Y's in the Within-Cell Correlations Analysis report lets you determine if multicollinearity is a problem. If this R^2 value is greater than .99 for any variable, you should take corrective action (remove one of the variables). To correct for multicollinearity, begin removing the variables one at a time until all of the R^2's are less than .99. Do not remove them all at once! *Singularity* is the extreme form of multicollinearity in which the R^2 value is one.

Forms of multicollinearity may show up when you have very small cell sample sizes (when the number of observations is less than the number of variables). In this case, you must reduce the number of dependent variables.

Data Structure

The data must be entered in a format that places the dependent variables and values of each factor side by side. An example of the data for a MANOVA design is shown in the table below. In this example, *WRATR* and *WRATA* are the two dependent variables. *Treatment* and *Disability* are two factor variables. This database is stored in the file MANOVA1.

MANOVA1 Dataset (Subset)

WRATR	WRATA	Treatment	Disability
115	108	1	1
98	105	1	1
107	98	1	1
90	92	2	1
85	95	2	1
80	81	2	1
100	105	1	2
105	95	1	2
95	98	1	2
70	80	2	2
.	.	.	.
.	.	.	.
.	.	.	.

Unequal Sample Size and Missing Data

You should begin by screening your data. Pay particular attention to patterns of missing values. When using MANOVA, you should have more observations per factor category than you have dependent variables so that you can test the equality of covariance matrices using Box's M test.

NCSS ignores rows with missing values. If it appears that most of the missing values occur in one or two variables, you might want to leave these out of the analysis in order to obtain more data and hence more power.

NCSS uses the GLM procedure for calculating the hypothesis and error matrices. Each matrix is calculated as if it were fit last in the model. This is the recommended way of obtaining these matrices. This method is valid even when the sample sizes for the various groups are unequal.

Example 1 – Multivariate Analysis of Variance

This section presents an example of how to run an analysis of the data contained in the MANOVA1 dataset.

Setup

To run this example, complete the following steps:

1 Open the MANOVA1 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **MANOVA1** and click **OK**.

2 Specify the Multivariate Analysis of Variance (MANOVA) procedure options

- Find and open the **Multivariate Analysis of Variance (MANOVA)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab

Response Variables..**WRATR-WRATA**
Factor Variable 1 ....**Treatment**
Factor Variable 2 ....**Disability**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Expected Mean Squares

```
Expected Mean Squares

Source                     Term       Denominator    Expected
Term              DF       Fixed?     Term           Mean Square
A: Treatment      1        Yes        S(AB)          S+bsA
B: Disability     2        Yes        S(AB)          S+asB
AB                2        Yes        S(AB)          S+sAB
S(AB)             12       No                        S

Note: Expected Mean Squares are for the balanced cell-frequency case.
```

The Expected Mean Square expressions are provided to show the appropriate error term for each factor. The correct error term for a factor is that term that is identical except for the factor being tested.

Source Term

The source of variation or term in the model.

DF

The degrees of freedom. The number of observations “used” by this term.

Term Fixed?

Indicates whether the term is “fixed” or “random.”

Denominator Term

Indicates the term used as the denominator in the F-ratio.

Expected Mean Square

This expression represents the expected value of the corresponding mean square if the design was completely balanced. “S” represents the expected value of the mean square error (sigma). The uppercase letters represent either the adjusted sum of squared treatment means if the factor is fixed, or the variance component if the factor is random. The lowercase letter represents the number of levels for that factor, and “s” represents the number of replications of the experimental layout.

These EMS expressions are provided to determine the appropriate error term for each factor. The correct error term for a factor is that term whose EMS is identical except for the factor being tested.

MANOVA Tests

Term(DF): Name Test Statistic	Test Value	DF1	DF2	F-Ratio	P-Value	Decision (α = 0.05)
A(1): Treatment						
Wilks' Lambda	0.137721	2	11	34.44	0.000018	Reject
Lawley - Hotelling Trace	6.261036	2	11	34.44	0.000018	Reject
Pillai's Trace	0.862279	2	11	34.44	0.000018	Reject
Roy's Largest Root	6.261036	2	11	34.44	0.000018	Reject
F (WRATR)	2090.888889	1	12	46.12	0.000019	Reject
F (WRATA)	1494.222222	1	12	33.25	0.000089	Reject
B(2): Disability						
Wilks' Lambda	0.255263	4	22	5.39	0.003528	Reject
Lawley - Hotelling Trace	2.895034	4	20	7.24	0.000896	Reject
Pillai's Trace	0.750481	4	24	3.60	0.019460	Reject
Roy's Largest Root	2.887241	2	12	17.32	0.000290	Reject
F (WRATR)	260.388889	2	12	5.74	0.017784	Reject
F (WRATA)	563.388889	2	12	12.54	0.001151	Reject
AB(2)						
Wilks' Lambda	0.908068	4	22	0.27	0.893037	Accept
Lawley - Hotelling Trace	0.100954	4	20	0.25	0.904790	Accept
Pillai's Trace	0.092192	4	24	0.29	0.881598	Accept
Roy's Largest Root	0.098039	2	12	0.59	0.570550	Accept
F (WRATR)	1.055556	2	12	0.02	0.977029	Accept
F (WRATA)	26.388889	2	12	0.59	0.571116	Accept

This report gives the results of the various significance tests. Usually, the four multivariate tests will lead to the same conclusions. When they do not, refer to the discussion of these tests found earlier in this chapter.

Once a multivariate test has found a term significant, use the univariate ANOVA to determine which of the variables and factors are “causing” the significance.

Term(DF): Name

The term in the design model with the degrees of freedom of the term in parentheses. For main effects, the variable name is also given.

Test Statistic

The name of the statistical test shown on this row of the report. The four multivariate tests are followed by the univariate F-tests of each variable.

Test Value

The value of the test statistic.

DF1

The numerator degrees of freedom of the F-ratio corresponding to this test.

DF2

The denominator degrees of freedom of the F-ratio corresponding to this test.

F-Ratio

The value of the F-test corresponding to this test. In some cases, this is an exact test. In other cases, this is an approximation to the exact test. See the discussion of each test to determine if it is exact or approximate.

P-Value

The significance level of the above F-ratio. The probability of an F-ratio larger than that obtained by this analysis. For example, to test at an alpha of 0.05, this probability would have to be less than 0.05 to make the F-ratio significant.

Decision (α = 0.05)

The decision to accept or reject the null hypothesis at the given level of significance. Note that you specify the level of significance when you select Alpha.

Correlations and Within-Group Covariances

Correlations and Within-Group Covariances

	WRATR	WRATA
WRATR	45.33333	2.583333
WRATA	0.0572313	44.94444

The correlations are shown in the off-diagonals in the lower-left half of the matrix.
The within-group covariances are shown on the diagonal and in the off-diagonals in the upper-right half of the matrix.

This report displays the correlations and covariances formed by averaging across all of the individual group covariance matrices. The correlations are shown in the lower-left half of the matrix. The within-group covariances are shown on the diagonal and in the upper-right half of the matrix.

Within-Cell Correlations Analysis

Within-Cell Correlations Analysis

Variable	R-Squared Other Y's	Canonical Variate	Eigenvalue	Percent of Total	Cumulative Total
WRATR	0.003275	1	1.057231	52.86	52.86
WRATA	0.003275	2	0.942769	47.14	100.00

This report analyzes the within-cell correlation matrix. It lets you diagnose multicollinearity problems as well as determine the number of dimensions that are being used. This is useful in determining if Pillai's trace should be used.

R-Squared Other Y's

This is the R-Squared index of this variable with the other variables. When this value is larger than 0.99, severe multicollinearity problems exist. If this happens, you should remove the variable with the largest R-Squared and re-run your analysis.

Canonical Variate

The identification numbers of the canonical variates that are generated during the analysis. The total number of variates is the smaller of the number of variables and the number of degrees of freedom in the model.

Eigenvalue

The eigenvalues of the within correlation matrix. Note that this value is not associated with the variable at the beginning of the row, but rather with the canonical variate number directly to the left.

Percent of Total

The percent that the eigenvalue is of the total. Note that the sum of the eigenvalues will equal the number of variates. If the percentage accounted for by the first eigenvalue is relatively large (70 or 80 percent), Pillai's trace will be less powerful than the other three multivariate tests.

Cumulative Total

The cumulative total of the Percent of Total column.

Univariate Analysis of Variance (ANOVA) Tables

Analysis of Variance (ANOVA) Table for WRATR

Source Term	DF	Sum of Squares	Mean Mean Square	F-Ratio	P-Value	Power (α = 0.05)
A: Treatment	1	2090.889	2090.889	46.12	0.000019*	0.999988
B: Disability	2	520.7778	260.3889	5.74	0.017784*	0.763859
AB	2	2.111111	1.055556	0.02	0.977029	0.052757
S	12	544	45.33333			
Total (Adjusted)	17	3157.778				
Total	18					

* Term significant at alpha = 0.05

Analysis of Variance (ANOVA) Table for WRATA

Source Term	DF	Sum of Squares	Mean Mean Square	F-Ratio	P-Value	Power (α = 0.05)
A: Treatment	1	1494.222	1494.222	33.25	0.000089*	0.999519
B: Disability	2	1126.778	563.3889	12.54	0.001151*	0.981144
AB	2	52.77778	26.38889	0.59	0.571116	0.125682
S	12	539.3333	44.94444			
Total (Adjusted)	17	3213.111				
Total	18					

* Term significant at alpha = 0.05

This is the standard ANOVA report as documented in the General Linear Models chapter. A separate report is displayed for each of the response variables.

Means and Standard Errors

Means and Standard Errors of WRATR

Term	Count	Mean	Standard Error*
All	18	89.11111	
A: Treatment			
1	9	99.88889	2.234687
2	9	78.33334	2.234687
B: Disability			
1	6	95.83334	2.736922
2	6	88.83334	2.736922
3	6	82.66666	2.736922
AB: Treatment, Disability			
1, 1	3	106.6667	3.870592
1, 2	3	100	3.870592
1, 3	3	93	3.870592
2, 1	3	85	3.870592
2, 2	3	77.66666	3.870592
2, 3	3	72.33334	3.870592

* The standard errors are calculated using the mean square error from the corresponding Analysis of Variance (ANOVA) table.

Means and Standard Errors of WRATA

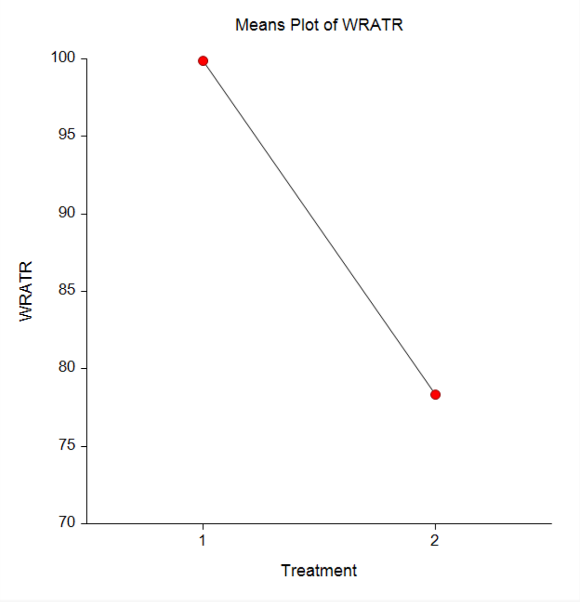
Term	Count	Mean	Standard Error*
All	18	87.22222	
A: Treatment			
1	9	96.33334	2.234687
2	9	78.11111	2.234687
B: Disability			
1	6	96.5	2.736922
2	6	88	2.736922
3	6	77.16666	2.736922
AB: Treatment, Disability			
1, 1	3	103.6667	3.870592
1, 2	3	99.33334	3.870592
1, 3	3	86	3.870592
2, 1	3	89.33334	3.870592
2, 2	3	76.66666	3.870592
2, 3	3	68.33334	3.870592

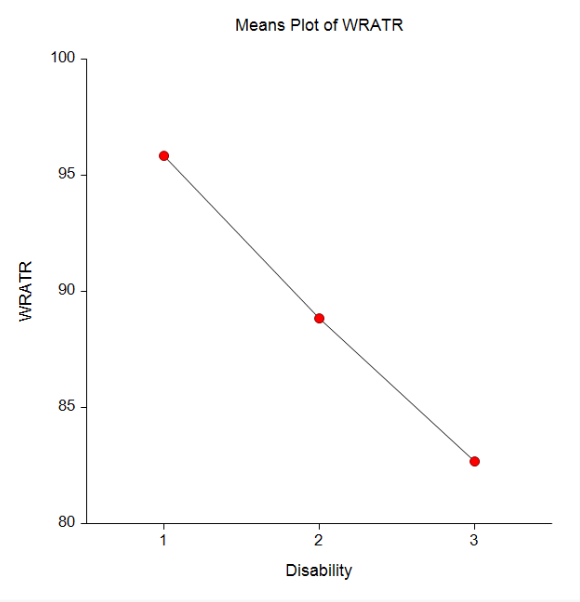
* The standard errors are calculated using the mean square error from the corresponding Analysis of Variance (ANOVA) table.

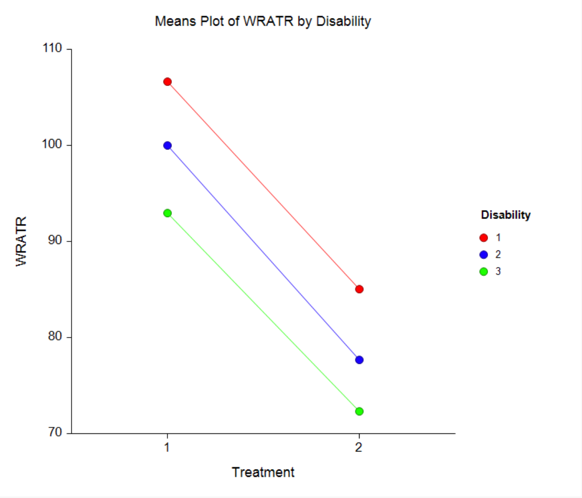
This report provides the least-squares means and standard errors for each variable. Note that the standard errors are calculated using the mean square error from the ANOVA table. They are not the standard errors that would be calculated from the individual cells.

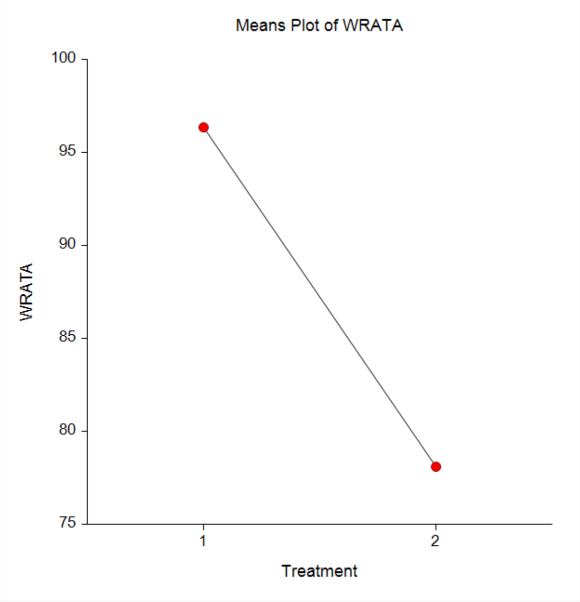
Means Plots

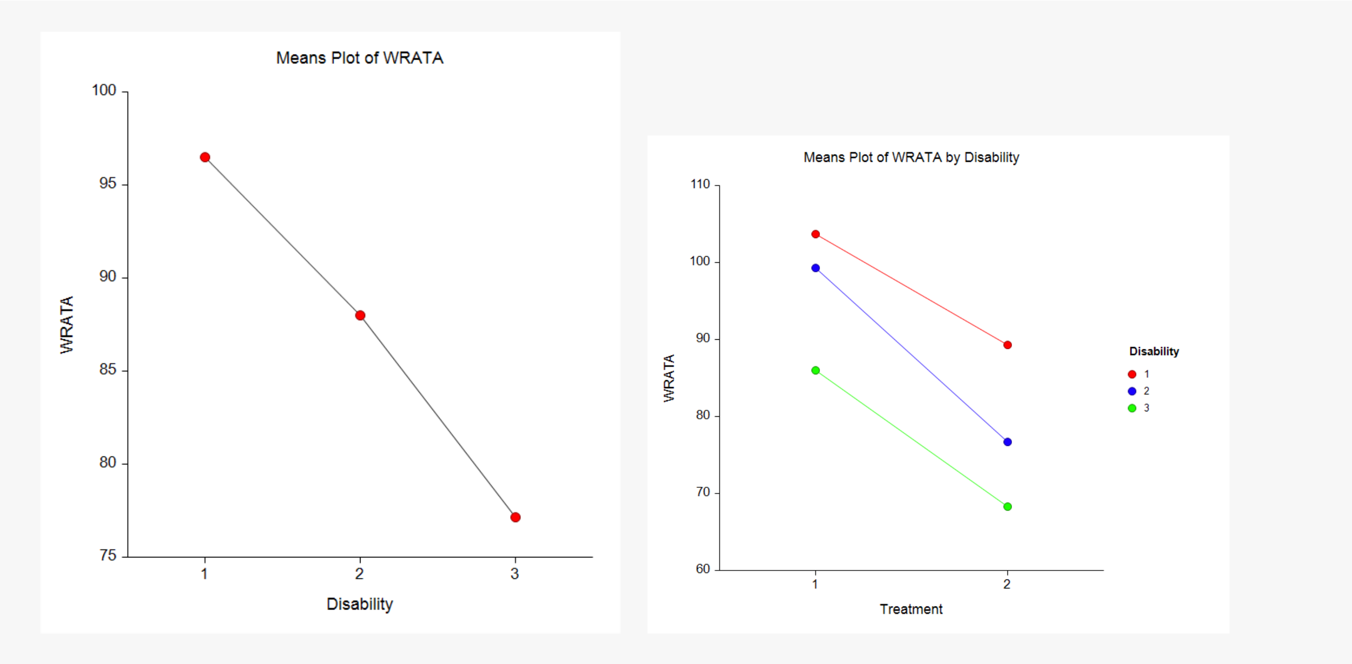
Means Plots











This section provides the means plots for main effects and interactions.