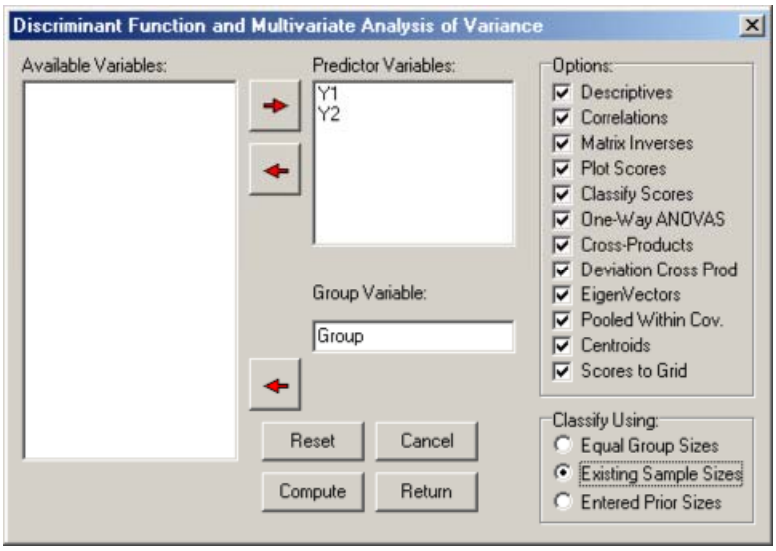
Discriminant Function / MANOVA

An Example

We will use the file labeled ManoDiscrim.LAZ for our example. A file of the same name (or a .tab file) should be in your directory. Load the file and then click on the Analyses / Multivariate / Discriminant Function option. You should see the form below completed for a discriminant function analysis:



Discriminant Function Analysis Form

You will notice we have asked for all options and have specified that classification use the a priori (sample) sizes for classification. When you click the Compute button, the following results are obtained:

```
MULTIVARIATE ANOVA / DISCRIMINANT FUNCTION
Reference: Multiple Regression in Behavioral Research
Elazar J. Pedhazur, 1997, Chapters 20-21
Harcourt Brace College Publishers

Total Cases := 15, Number of Groups := 3

SUM OF CROSS-PRODUCTS forGroup 1, N = 5 with    5 valid cases.

Variables
                    Y1           Y2
       Y1      111.000      194.000
       Y2      194.000      343.000


WITHIN GROUP SUM OF DEVIATION CROSS-PROD with    5 valid cases.
```

Variables

	Y1	Y2
Y1	5.200	5.400
Y2	5.400	6.800

MEANS FOR GROUP 1, N := 5 with 5 valid cases.

Variables	Y1	Y2
	4.600	8.200

VARIANCES FOR GROUP 1 with 5 valid cases.

Variables	Y1	Y2
	1.300	1.700

STANDARD DEVIATIONS FOR GROUP 1 with 5 valid cases.

Variables	Y1	Y2
	1.140	1.304

SUM OF CROSS-PRODUCTS forGroup 2, N = 5 with 5 valid cases.

Variables

	Y1	Y2
Y1	129.000	169.000
Y2	169.000	223.000

WITHIN GROUP SUM OF DEVIATION CROSS-PROD with 5 valid cases.

Variables

	Y1	Y2
Y1	4.000	4.000
Y2	4.000	5.200

MEANS FOR GROUP 2, N := 5 with 5 valid cases.

Variables	Y1	Y2
	5.000	6.600

VARIANCES FOR GROUP 2 with 5 valid cases.

Variables	Y1	Y2
	1.000	1.300

STANDARD DEVIATIONS FOR GROUP 2 with 5 valid cases.

Variables	Y1	Y2
	1.000	1.140

SUM OF CROSS-PRODUCTS forGroup 3, N = 5 with 5 valid cases.

Variables

	Y1	Y2
Y1	195.000	196.000
Y2	196.000	199.000

WITHIN GROUP SUM OF DEVIATION CROSS-PROD with 5 valid cases.

Variables

	Y1	Y2
Y1	2.800	3.800
Y2	3.800	6.800

MEANS FOR GROUP 3, N := 5 with 5 valid cases.

Variables	Y1	Y2
	6.200	6.200

VARIANCES FOR GROUP 3 with 5 valid cases.

Variables	Y1	Y2
	0.700	1.700

STANDARD DEVIATIONS FOR GROUP 3 with 5 valid cases.

Variables	Y1	Y2
	0.837	1.304

TOTAL SUM OF CROSS-PRODUCTS with 15 valid cases.

Variables

	Y1	Y2
Y1	435.000	559.000
Y2	559.000	765.000

TOTAL SUM OF DEVIATION CROSS-PRODUCTS with 15 valid cases.

Variables

	Y1	Y2
Y1	18.933	6.000
Y2	6.000	30.000

MEANS with 15 valid cases.

Variables	Y1	Y2
	5.267	7.000

VARIANCES with 15 valid cases.

Variables	Y1	Y2
	1.352	2.143

STANDARD DEVIATIONS with 15 valid cases.

Variables	Y1	Y2
	1.163	1.464

BETWEEN GROUPS SUM OF DEV. CPs with 15 valid cases.

Variables

	Y1	Y2
Y1	6.933	-7.200
Y2	-7.200	11.200

UNIVARIATE ANOVA FOR VARIABLE Y1

SOURCE	DF	SS	MS	F	PROB > F
BETWEEN	2	6.933	3.467	3.467	0.065
ERROR	12	12.000	1.000		
TOTAL	14	18.933			

UNIVARIATE ANOVA FOR VARIABLE Y2

SOURCE	DF	SS	MS	F	PROB > F

BETWEEN	2	11.200	5.600	3.574	0.061
ERROR	12	18.800	1.567		
TOTAL	14	30.000			

Inv. of Pooled Within Dev. CPs Matrix with 15 valid cases.
Variables

	Y1	Y2
Y1	0.366	-0.257
Y2	-0.257	0.234

Number of roots extracted := 2
Percent of trace extracted := 100.0000
Roots of the W inverse time B Matrix

No.	Root	Proportion	Canonical R	Chi-Squared	D.F.	Prob.
1	8.7985	0.9935	0.9476	25.7156	4	0.000
2	0.0571	0.0065	0.2325	0.6111	1	0.434

Eigenvectors of the W inverse x B Matrix with 15 valid cases.
Variables

	1	2
Y1	-2.316	0.188
Y2	1.853	0.148

Pooled Within-Groups Covariance Matrix with 15 valid cases.
Variables

	Y1	Y2
Y1	1.000	1.100
Y2	1.100	1.567

Total Covariance Matrix with 15 valid cases.
Variables

	Y1	Y2
Y1	1.352	0.429
Y2	0.429	2.143

Raw Function Coeff.s from Pooled Cov. with 15 valid cases.
Variables

	1	2
Y1	-2.030	0.520
Y2	1.624	0.409

Raw Discriminant Function Constants with 15 valid cases.

Variables	1	2
	-0.674	-5.601

Fisher Discriminant Functions
Group 1 Constant := -24.402

Variable	Coefficient
1	-5.084
2	8.804

Group 2 Constant := -14.196

Variable	Coefficient
1	1.607
2	3.084

Group 3 Constant := -19.759

Variable	Coefficient
1	8.112
2	-1.738

CLASSIFICATION OF CASES

SUBJECT ID NO.	ACTUAL GROUP	HIGH IN	PROBABILITY GROUP P(G/D)	SEC.D GROUP	HIGH P(G/D)	DISCRIM SCORE
1	1	1	0.9999	2	0.0001	4.6019
						-1.1792
2	1	1	0.9554	2	0.0446	2.5716
						-0.6590
3	1	1	0.8903	2	0.1097	2.1652
						0.2699
4	1	1	0.9996	2	0.0004	3.7890
						0.6786
5	1	1	0.9989	2	0.0011	3.3826
						1.6075
6	2	2	0.9746	3	0.0252	-0.6760
						-1.4763
7	2	2	0.9341	1	0.0657	0.9478
						-1.0676
8	2	2	0.9730	1	0.0259	0.5414
						-0.1387
9	2	2	0.5724	3	0.4276	-1.4888
						0.3815
10	2	2	0.9842	1	0.0099	0.1350
						0.7902
11	3	3	0.9452	2	0.0548	-2.7062
						-0.9560
12	3	3	0.9999	2	0.0001	-4.7365
						-0.4358
13	3	3	0.9893	2	0.0107	-3.1126
						-0.0271
14	3	3	0.9980	2	0.0020	-3.5191
						0.9018
15	3	3	0.8007	2	0.1993	-1.8953
						1.3104

CLASSIFICATION TABLE

Variables	PREDICTED GROUP 1	2	3	TOTAL
1	5	0	0	5
2	0	5	0	5
3	0	0	5	5
TOTAL	5	5	5	15

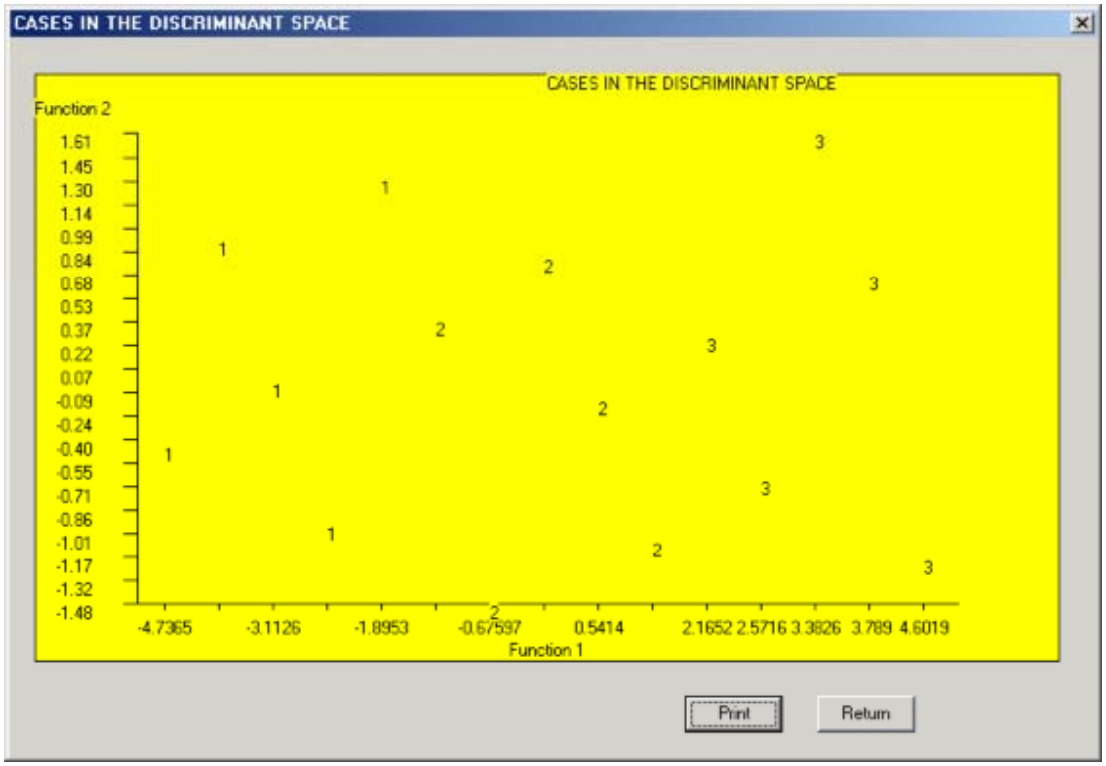
Standardized Coeff. from Pooled Cov. with 15 valid cases.

Variables	1	2
Y1	-2.030	0.520
Y2	2.032	0.511

Centroids with 15 valid cases.

Variables	1	2
1	3.302	0.144

```
          2      -0.108       -0.302
          3      -3.194        0.159

Raw Coefficients from Total Cov. with   15 valid cases.
Variables
                       1            2
       Y1       -0.701        0.547
       Y2        0.560        0.429

Raw Discriminant Function Constants with   15 valid cases.
Variables            1            2
                -0.674       -5.601


Standardized Coeff.s from Total Cov. with   15 valid cases.
Variables
                       1            2
       Y1       -0.815        0.636
       Y2        0.820        0.628

Total Correlation Matrix with   15 valid cases.
Variables
                     Y1           Y2
       Y1         1.000        0.252
       Y2         0.252        1.000

Corr.s Between Variables and Functions with   15 valid cases.
Variables
                       1            2
       Y1       -0.608        0.794
       Y2        0.615        0.788

Wilk's Lambda =     0.0965.
F =    12.2013 with D.F.     4 and    22 .  Prob > F = 0.0000

Bartlett Chi-Squared =    26.8845 with 4 D.F. and prob. = 0.0000
```



Plot of Cases in a Discriminant Space

```
Pillai Trace =     0.9520
```

You will notice that we have obtained cross-products and deviation cross-products for each group as well as the combined between and within groups as well as descriptive statistics (means, variances, standard deviations.) Two roots were obtained, the first significant at the 0.05 level using a chi-square test. The one-way analyses of variances completed for each continuous variable were not significant at the 0.05 level which demonstrates that a multivariate analysis may identify group differences not caught by individual variable analysis. The discriminant functions can be used to plot the group subjects in the (orthogonal) space of the functions. If you examine the plot you can see that the individuals in the three groups analyzed are easily separated using just the first discriminant function (the horizontal axis.) Raw and standardized coefficients for the discriminant functions are presented as well as Fisher's discriminant functions for each group. The latter are used to classify the subjects and the classifications are shown along with a table which summarizes the classifications. Note that in this example, all cases are correctly classified. Certainly, a cross-validation of the functions for classification would likely encounter some errors of classification. Since we asked that the discriminant scores be placed in the data grid, the last figure shows the data grid with the Fisher discriminant scores saved as two new variables.