What are some examples of One-way MANOVA data analysis using SAS?

Authored by
stats writer

June 29, 2024

RECOMMENDED CITATION

stats writer (2024). *What are some examples of One-way MANOVA data analysis using SAS?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=158686

The one-way MANOVA (multivariate analysis of variance) is a statistical method used to analyze the differences between three or more groups on multiple dependent variables simultaneously. This type of analysis is commonly used in research settings to determine if there are significant differences between groups on a set of related variables. SAS (Statistical Analysis System) is a software program commonly used for data analysis and has several functions available for conducting one-way MANOVA. Some examples of one-way MANOVA analyses that can be performed using SAS include comparing the average scores of students from three different schools on multiple academic subjects, investigating the effects of different treatments on several health outcomes in a clinical trial, or examining the impact of different marketing strategies on various consumer behaviors. Overall, one-way MANOVA analysis using SAS provides a comprehensive approach for analyzing data from multiple groups and can be applied in a variety of research fields.

One-way MANOVA | SAS Data Analysis Examples

Version info: Code for this page was tested in SAS 9.3

MANOVA is used to model two or more dependent variables that are continuous with one or more categorical predictor variables.

Please note: The purpose of this page is to show how to use various data analysis commands. It does not cover all aspects of the research process which researchers are expected to do. In particular, it does not cover data cleaning and checking, verification of assumptions,

model diagnostics or
potential follow-up analyses.

Examples of one-way multivariate analysis of variance

Example 1. A researcher randomly assigns 33 subjects to one of three groups. The first group receives technical dietary information interactively from an on-line website. Group 2 receives the same information from a nurse practitioner, while group 3 receives the information from a video tape made by the same nurse practitioner. The researcher looks at three different ratings of the presentation, difficulty, usefulness and importance, to determine if there is a difference in the modes of presentation. In particular, the researcher is interested in whether the interactive website is superior because that is the most cost-effective way of delivering the information.

Example 2. A clinical psychologist recruits 100 people who suffer from panic disorder into his study. Each subject receives one

of four types of
treatment for eight weeks. At the end of treatment, each subject
participates in a structured interview, during which the clinical psychologist
makes three ratings: physiological, emotional and cognitive. The
clinical psychologist wants to know which type of treatment most reduces the
symptoms of the panic disorder as measured on the physiological, emotional and
cognitive scales. (This example was adapted from Grimm and Yarnold, 1995,
page 246.)

Description of the data

Let's pursue Example 1 from above.

We have a data file, manova,
with 33 observations on three response variables.
The response variables are ratings of useful, difficulty and importance.
Level 1 of the group variable is the treatment group, level 2 is control group 1 and

level 3 is control group 2.

Let's look at the data. It is always a good idea to start with descriptive statistics.

```
proc means data = mylib.manova;
var difficulty useful importance;
run;
```

The MEANS Procedure

Variable N Mean Std Dev Minimum Maximum

DIFFICULTY 33 5.7151515 2.0175978 2.4000001 10.2500000

USEFUL 33 16.3303030 3.2924615 11.8999996 24.2999992

IMPORTANCE 33 6.4757576 3.9851309 0.2000000 18.7999992

```
proc freq data = mylib.manova;
tables group;
run;
```

The FREQ Procedure

Cumulative Cumulative
GROUP Frequency Percent Frequency Percent
---
1 11 33.33 11 33.33
2 11 33.33 22 66.67
3 11 33.33 33 100.00

```
proc means n mean std min max data = mylib.manova;
class group;
var useful difficulty importance;
run;
```

The MEANS Procedure

N
GROUP Obs Variable N Mean Std Dev Minimum Maximum
---
1 11 USEFUL 11 18.1181817 3.9037974 13.0000000 24.2999992
DIFFICULTY 11 6.1909091 1.8997129 3.7500000 10.2500000

IMPORTANCE 11 8.6818181 4.8630890 3.3000000 18.7999992

2 11 USEFUL 11 15.5272729 2.0756162 12.8000002 19.7000008
DIFFICULTY 11 5.5818183 2.4342631 2.4000001 9.8500004
IMPORTANCE 11 5.1090909 2.5311873 0.2000000 8.5000000

3 11 USEFUL 11 15.3454545 3.1382682 11.8999996 19.7999992
DIFFICULTY 11 5.3727273 1.7590287 2.6500001 8.7500000
IMPORTANCE 11 5.6363637 3.5469065 0.7000000 10.3000002
--

```
proc corr data = mylib.manova nosimple;
var useful difficulty importance;
run;
```

The CORR Procedure

3 Variables: USEFUL DIFFICULTY IMPORTANCE

Pearson Correlation Coefficients, N = 33
Prob > |r| under H0: Rho=0

USEFUL DIFFICULTY IMPORTANCE

USEFUL 1.00000 0.09783 -0.34112
0.5881 0.0520

DIFFICULTY 0.09783 1.00000 0.19782
0.5881 0.2698

IMPORTANCE -0.34112 0.19782 1.00000
0.0520 0.2698

Analysis methods you might consider

Below is a list of some analysis methods you may have encountered. Some of the methods listed are quite reasonable, while others
have either fallen out of favor or have limitations.

One-way MANOVA

We will use proc glm to run the one-way MANOVA. We will list the

variable group on the class statement to indicate that it is a
categorical predictor variable. We use the ss3 option on the
model statement to get only the Type III sums of squares in the output.
We use some contrast statements to specify two contrasts in which we are
interested. We will discuss these when we see their output. We use
the first manova statement to obtain all of the multivariate tests that
SAS offers; we use the second manova statement to run the multivariate
tests using only the variables useful and importance.

Because the output is very long, we will break it up and discuss the
different sections individually. Please also see our
Annotated Output: SAS MANOVA.

```
proc glm data= mylib.manova;
class group;
model useful difficulty importance = group / ss3;
```

contrast '1 vs 2&3' group 2 -1 -1;
contrast '2 vs 3' group 0 1 -1;
manova h=_all_;
manova h=group m=(1 0 1);
run;
The GLM Procedure

Class Level Information

Class Levels Values

GROUP 3 1 2 3

Number of Observations Read 33
Number of Observations Used 33

Dependent Variable: USEFUL

Sum of
Source DF Squares Mean Square F Value Pr > F

Model 2 52.9242378 26.4621189 2.70 0.0835

Error 30 293.9654425 9.7988481

Corrected Total 32 346.8896803

R-Square Coeff Var Root MSE USEFUL Mean

0.152568 19.16873 3.130311 16.33030

Source DF Type III SS Mean Square F Value Pr > F

GROUP 2 52.92423783 26.46211891 2.70 0.0835

Contrast DF Contrast SS Mean Square F Value Pr > F

1 vs 2&3 1 52.74241913 52.74241913 5.38 0.0273
2 vs 3 1 0.18181870 0.18181870 0.02 0.8926

Dependent Variable: DIFFICULTY

Sum of
Source DF Squares Mean Square F Value Pr > F

Model 2 3.9751512 1.9875756 0.47 0.6282

Error 30 126.2872767 4.2095759

Corrected Total 32 130.2624279

R-Square Coeff Var Root MSE DIFFICULTY Mean

0.030516 35.89975 2.051725 5.715152

Source DF Type III SS Mean Square F Value Pr > F

GROUP 2 3.97515121 1.98757560 0.47 0.6282

Contrast DF Contrast SS Mean Square F Value Pr > F

1 vs 2&3 1 3.73469643 3.73469643 0.89 0.3538
2 vs 3 1 0.24045478 0.24045478 0.06 0.8127

Dependent Variable: IMPORTANCE

Sum of
Source DF Squares Mean Square F Value Pr > F

Model 2 81.8296936 40.9148468 2.88 0.0718

Error 30 426.3708962 14.2123632

Corrected Total 32 508.2005898

R-Square Coeff Var Root MSE IMPORTANCE Mean

0.161018 58.21603 3.769929 6.475758

Source DF Type III SS Mean Square F Value Pr > F

GROUP 2 81.82969356 40.91484678 2.88 0.0718

Contrast DF Contrast SS Mean Square F Value Pr > F

1 vs 2&3 1 80.30060224 80.30060224 5.65 0.0240
2 vs 3 1 1.52909132 1.52909132 0.11 0.7452

Next, we will look at the overall MANOVA itself.

Multivariate Analysis of Variance

Characteristic Roots and Vectors of: E Inverse * H, where
H = Type III SSCP Matrix for GROUP
E = Error SSCP Matrix

Characteristic Characteristic Vector V'EV=1
Root Percent USEFUL DIFFICULTY IMPORTANCE

0.89198790 99.42 0.06410227 -0.00186162 0.05375069
0.00524207 0.58 0.01442655 0.06888878 -0.02620577
0.00000000 0.00 -0.03149580 0.05943387 0.01270798

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall GROUP Effect
H = Type III SSCP Matrix for GROUP
E = Error SSCP Matrix

S=2 M=0 N=13

Statistic Value F Value Num DF Den DF Pr > F

Wilks' Lambda 0.52578838 3.54 6 56 0.0049
Pillai's Trace 0.47667013 3.02 6 58 0.0122
Hotelling-Lawley Trace 0.89722998 4.12 6 35.61 0.0031
Roy's Greatest Root 0.89198790 8.62 3 29 0.0003

NOTE: F Statistic for Roy's Greatest Root is an upper bound.
NOTE: F Statistic for Wilks' Lambda is exact.

Characteristic Roots and Vectors of: E Inverse * H, where
H = Contrast SSCP Matrix for 1 vs 2&3
E = Error SSCP Matrix

Characteristic Characteristic Vector V'EV=1
Root Percent USEFUL DIFFICULTY IMPORTANCE

0.89039367 100.00 0.06414887 -0.00163749 0.05366515
0.00000000 0.00 -0.01449686 0.09003145 -0.00766730
0.00000000 0.00 0.03136839 0.01315947 -0.02826015

The overall multivariate test is significant, which means

that differences
between the levels of the variable group exist. To find where the
differences lie, we will follow up with several post-hoc tests. We will begin with the multivariate test of group 1 versus the
average of groups 2 and 3.

/* contrast '1 vs 2&3' group 2 -1 -1; manova h-_all_; */

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall 1 vs 2&3 Effect
H = Contrast SSCP Matrix for 1 vs 2&3
E = Error SSCP Matrix

S=1 M=0.5 N=13

Statistic Value F Value Num DF Den DF Pr > F

Wilks' Lambda 0.52899035 8.31 3 28 0.0004
Pillai's Trace 0.47100965 8.31 3 28 0.0004
Hotelling-Lawley Trace 0.89039367 8.31 3 28 0.0004
Roy's Greatest Root 0.89039367 8.31 3 28 0.0004

Taking all three dependent variables together, this

contrast is statistically significant.

Here is the multivariate test of group 2 versus group 3.

/* contrast '2 vs 3' group 0 1 -1; manova h-_all_; */

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall 2 vs 3 Effect
H = Contrast SSCP Matrix for 2 vs 3
E = Error SSCP Matrix

S=1 M=0.5 N=13

Statistic Value F Value Num DF Den DF Pr > F

Wilks' Lambda 0.99321011 0.06 3 28 0.9785
Pillai's Trace 0.00678989 0.06 3 28 0.9785
Hotelling-Lawley Trace 0.00683631 0.06 3 28 0.9785
Roy's Greatest Root 0.00683631 0.06 3 28 0.9785

Taking all three dependent variables together, this contrast is not statistically significant.

We know from the univariate tests above that difficulty

by itself was clearly not significant. This next test does the multivariate test using the combination of useful and importance.

/* manova h=group m=(1 0 1); */

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall GROUP Effect
on the Variables Defined by the M Matrix Transformation
H = Type III SSCP Matrix for GROUP
E = Error SSCP Matrix

S=1 M=0 N=14

Statistic Value F Value Num DF Den DF Pr > F

Wilks' Lambda 0.53598494 12.99 2 30

The multivariate test with useful and importance as dependent variables and group as the independent variable is statistically significant.

We can use the lsmeans statement to obtain adjusted

predicted values for each of the dependent variables for each of the groups. These values can be helpful in seeing where differences between levels of the predictor variable are and describing the model.

****** STOP HERE AND REVIEW ******

```
proc glm data= mylib.manova;
class group;
model useful difficulty importance = group / ss3;
lsmeans group;
run;
```

<SOME OUTPUT OMITTED**>**

The GLM Procedure

Least Squares Means

USEFUL

GROUP	LSMEAN
1	18.1181817
2	15.5272729
3	15.3454545

DIFFICULTY

GROUP LSMEAN

1 6.19090908

2 5.58181828

3 5.37272726

IMPORTANCE

GROUP LSMEAN

1 8.68181812

2 5.10909089

3 5.63636369

In each of the three columns above, we see that the predicted means for groups 2 and 3 are very similar; the predicted mean for group 1 is higher than those for groups 2 and 3.

In the examples below, we obtain the differences in the means for each of the dependent variables for each of the control groups (groups 2 and 3) compared to the treatment group (group1), by specifying group 1 to

be the reference group (called "control" by SAS, confusingly for this scenario). With respect to the dependent variable useful, the difference between the means for control group 1 versus the treatment group is approximately -2.59 (15.53 - 18.12). The difference between the means for control group 2 versus the treatment group is approximately -2.77 (15.35 - 18.12). With respect to the dependent variable difficulty, the difference between the means for control group 1 versus the treatment group is approximately -0.61 (5.58 - 6.19). The difference between the means for control group 2 versus the treatment group is approximately -0.82 (5.37 - 6.19).


```
proc glm data= mylib.manova;
class group;
model useful difficulty importance = group / ss3;
lsmeans group / pdiff = control('1') cl;
run;
```

The GLM Procedure

Least Squares Means

Adjustment for Multiple Comparisons: Dunnett

H0:LSMean=

USEFUL Control

GROUP LSMEAN Pr > |t|

1 18.1181817

2 15.5272729 0.1099

3 15.3454545 0.0836

USEFUL

GROUP LSMEAN 95% Confidence Limits

1 18.118182 16.190635 20.045728

2 15.527273 13.599726 17.454819

3 15.345454 13.417908 17.273001

Least Squares Means for Effect GROUP

Difference Simultaneous 95%

Between Confidence Limits for

i j Means LSMean(i)-LSMean(j)

2 1 -2.590909 -5.688577 0.506759

3 1 -2.772727 -5.870395 0.324941

H0:LSMean=
DIFFICULTY Control
GROUP LSMEAN Pr > |t|

1 6.19090908
2 5.58181828 0.7117
3 5.37272726 0.5518

DIFFICULTY
GROUP LSMEAN 95% Confidence Limits

1 6.190909 4.927522 7.454296
2 5.581818 4.318431 6.845206
3 5.372727 4.109340 6.636115

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Dunnett

Least Squares Means for Effect GROUP

Difference Simultaneous 95%
Between Confidence Limits for
i j Means LSMean(i)-LSMean(j)

2 1 -0.609091 -2.639420 1.421239

3 1 -0.818182 -2.848511 1.212148

H0:LSMean=

IMPORTANCE Control

GROUP LSMEAN Pr > |t|

1 8.68181812

2 5.10909089 0.0618

3 5.63636369 0.1203

IMPORTANCE

GROUP LSMEAN 95% Confidence Limits

1 8.681818 6.360415 11.003221

2 5.109091 2.787688 7.430494

3 5.636364 3.314961 7.957766

Least Squares Means for Effect GROUP

Difference Simultaneous 95%

Between Confidence Limits for

i j Means LSMean(i)-LSMean(j)

2 1 -3.572727 -7.303343 0.157889

3 1 -3.045454 -6.776070 0.685161

Finally, let's run separate univariate ANOVAs. Without a manova
statement specified, procglm will run separate ANOVAs when
multiple DVs are in the model statement.

```
proc glm data = mylib.manova;
class group;
model useful difficulty importance = group / ss3;
run;
```

Dependent Variable: USEFUL

Sum of
Source DF Squares Mean Square F Value Pr > F

Model 2 52.9242378 26.4621189 2.70 0.0835

Error 30 293.9654425 9.7988481

Corrected Total 32 346.8896803

R-Square Coeff Var Root MSE USEFUL Mean

0.152568 19.16873 3.130311 16.33030

Dependent Variable: DIFFICULTY

Sum of

Source DF Squares Mean Square F Value Pr > F

Model 2 3.9751512 1.9875756 0.47 0.6282

Error 30 126.2872767 4.2095759

Corrected Total 32 130.2624279

R-Square Coeff Var Root MSE DIFFICULTY Mean

0.030516 35.89975 2.051725 5.715152

Dependent Variable: IMPORTANCE

Sum of

Source DF Squares Mean Square F Value Pr > F

Model 2 81.8296936 40.9148468 2.88 0.0718

Error 30 426.3708962 14.2123632

Corrected Total 32 508.2005898

R-Square Coeff Var Root MSE IMPORTANCE Mean

0.161018 58.21603 3.769929 6.475758

None of the three ANOVAs were statistically significant at the alpha = .05 level.
In particular, the F-ratio for difficulty was less than 1.

Things to consider

See also

References