

Multiple Comparison (Post Hoc) Tests Matlab Tutorial

Assumptions (Same as ANOVA)

Data is independent and identically distributed (homogeneity of variance).
Data normally distributed.

How to load and format data

If you are unfamiliar with how to access MATLAB from your personal computer, look at the [Pratt Pundit](#). If you have not done so already, please review the instructions on ANOVA.

You can input your data into an Excel file. While you can have your data in rows or columns, here we will stick to data being laid out in columns.

For the drug data, your excel file would look like this:

	A	B	C	D	E	F
1						
2		Placebo	DrugOptA	DrugOptB	DrugOptC	DrugOptD
3		-1.12	-6.02	5.7	5.07	1.54
4		-0.98	-8.78	4.86	2.91	-0.79
5		-0.23	-7.48	5.95	-9.48	-5.6
6		-1.21	-9.4	3.79	-0.51	-2.53
7		-1.11	-8.35	3.57	3.66	-0.48
8		0.45	-8.02	4.54	-2.36	-6.07
9						
10						
11						
12						

You would have to upload this excel file into your [WebFiles](#).

You would then load the data into MATLAB using the command “xlsread,” like so:

```
>> data = xlsread('drugdata.xlsx')

data =

    -1.1200    -6.0200     5.7000     5.0700     1.5400
    -0.9800    -8.7800     4.8600     2.9100    -0.7900
    -0.2300    -7.4800     5.9500    -9.4800    -5.6000
    -1.2100    -9.4000     3.7900    -0.5100    -2.5300
    -1.1100    -8.3500     3.5700     3.6600    -0.4800
     0.4500    -8.0200     4.5400    -2.3600    -6.0700
```

You can separate your data into separate matrices by column in this fashion:

```
>> Placebo = data(:,1)
```

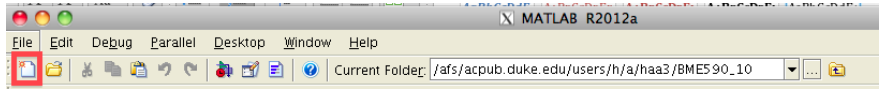
Placebo =

```
-1.1200
-0.9800
-0.2300
-1.2100
-1.1100
 0.4500
```

Quick Overview

These instructions on a multicompare test provided will require you to make a basic .m file. The text for the Matlab code can be found at the end of this document.

Click on the “new script” icon marked by a red box here:



The following will walk you through the commands you need to enter into your script. What literally needs to be written:

```
1 %Lines '3-6' are used to clear any stored data which may conflict with this
2 %program including overlapping variable names or not applicable data which
3 %could slow down run-time.
4
5 clc %Clear command screen
6 clf %Clear figures
7 clear all
8 close all hidden
9
10
11 placebo = [-1.12 -.98 .23 -1.21 -1.11 .45];
12 DrugOptA = [-6.02 -8.78 -7.48 -9.40 -8.35 -8.02];
13 DrugOptB = [5.70 4.86 5.95 3.79 3.57 4.54];
14 DrugOptC = [5.07 2.91 -9.48 -0.51 3.66 -2.36];
15 DrugOptD = [1.54 -0.79 -5.60 -2.53 -0.48 -6.07];
16
17
18
19 data(:,1) = placebo;
20 data(:,2) = DrugOptA';
21 data(:,3) = DrugOptB';
22 data(:,4) = DrugOptC';
23 data(:,5) = DrugOptD';
24
25
26 [p, table, stats] = anova1(data);
27 [c, m, h, nms] = multcompare(stats, 'alpha', .05, 'ctype', 'hsd');
28
```

Replace with your data here.

Replace with your data labels.

Replace with your tolerance value here

Replace with your type of test here. See "ctype values."

Ctype Values (from Mathworks)

Value	Description
'hsd' or 'tukey-kramer'	Use Tukey's honestly significant difference criterion. This is the default, and it is based on the Studentized range distribution. It is optimal for balanced one-way ANOVA and similar procedures with equal sample sizes. It has been proven to be conservative for one-way ANOVA with different sample sizes. According to the unproven Tukey-Kramer conjecture, it is also accurate for problems where the quantities being compared are correlated, as in analysis of covariance with unbalanced covariate values.
'lsd'	Use Tukey's least significant difference procedure. This procedure is a simple t-test. It is reasonable if the preliminary test (say, the one-way ANOVA <i>F</i> statistic) shows a significant difference. If it is used unconditionally, it provides no protection against multiple comparisons.
'bonferroni'	Use critical values from the t distribution, after a Bonferroni adjustment to compensate for multiple comparisons. This procedure is conservative, but usually less so than the Scheffé procedure.
'dunn-sidak'	Use critical values from the t distribution, after an adjustment for multiple comparisons that was proposed by Dunn and proved accurate by Sidák. This procedure is similar to, but less conservative than, the Bonferroni procedure.
'scheffe'	Use critical values from Scheffé's S procedure, derived from the F distribution. This procedure provides a simultaneous confidence level for comparisons of all linear combinations of the means, and it is conservative for comparisons of simple differences of pairs.

Detailed Instructions

Explanation of sample data:

The sample data tests the efficacy of a drug used to treat heart failure (HF). Five categories of data are shown, where each category is a type of drug (placebo vs. drug) or a dose of the drug. The dependent variable between the brackets is **change in HR**, which is a metric used to measure the efficacy of treatments for HF. If the code is run as is, $\alpha = 0.05$, and the Bonferroni correction is used. The results show that 'DrugOptA' causes the greatest decrease in HR, which is significant vs. placebo, while 'DrugOptB' is also significant but positively so. 'DrugOptC' and 'D' do not yield results that are significant from placebo.

Here's a breakdown of the script above:

Enter in your data. Notice here the data points for each matrix is separated by a space. If you'd like you can separate them by a comma instead.

```
10 %List of drugs (below) which are formatted with the name of the drug/drug
11 %option on the left followed by the dependent variable or results of the
12 %study on the right. Data entered this way are stored as 'n-by-1' vectors
13 - placebo = [-1.12 -.98 .23 -1.21 -1.11 .45];
14 - DrugOptA = [-6.02 -8.78 -7.48 -9.40 -8.35 -8.02];
15 - DrugOptB = [5.70 4.86 5.95 3.79 3.57 4.54];
16 - DrugOptC = [5.07 2.91 -9.48 -0.51 3.66 -2.36];
17 - DrugOptD = [1.54 -0.79 -5.60 -2.53 -0.48 -6.07];
18
```

Sort and organize your data for processing. This places your data in the format most friendly for MATLAB's 'anova1' and 'multcompare' commands.

```
20 %Combine the vectors above into a single array called 'data' where each row
21 %contains the result of a single drug or drug option
22 - data(:,1) = placebo;
23 - data(:,2) = DrugOptA';
24 - data(:,3) = DrugOptB';
25 - data(:,4) = DrugOptC';
26 - data(:,5) = DrugOptD';
27
```

Enter in the ANOVA and multcompare commands. For a more detailed description of the 'anova1' and 'multcompare' commands, visit the following Mathworks links: [anova1](#) and [multcompare](#).

```
29 - [p, table, stats] = anova1(data);
30 - [c, m, h, nms] = multcompare(stats, 'alpha', .05, 'ctype', 'bonferroni');
31
```

You'll notice these commands are for a Bonferroni test with a tolerance of 0.05. For another tolerance level, simply replace the 0.05 with your desired tolerance. For another multcompare test, like Tukey or Scheffe, replace the 'bonferroni', with your desired test as specified below in the Multcompare section

ANOVA1

Looking at line 29, we see:

```
[p, table, stats] = anova1(data);
```

where:

'p' returns the **p-value** under the null hypothesis that all samples stored in variable "data" are drawn from populations with the same mean.

'table' returns the **ANOVA table** shown below.

'stats' is the set of data which is used to perform a **follow-up multicomparison test**.

To review the relationship between ANOVA's and using a follow-up multicomparison test, the ANOVA test evaluates the hypothesis that the samples all have the same mean against the alternative that the means are not all the same. The multicomparison test is used to determine which pairs of means are significantly different, and which are not.

Multcompare

Looking at line 30, we see:

```
[c, m, h, nms] = multcompare(stats, 'alpha', .05, 'ctype', 'bonferroni');
```

INPUTS:

The 'multcompare' function is performed on the 'stats' structure which we generated using 'ANOVA1'

We can additionally set the alpha level using the 'alpha' function followed by our alpha-value (in this case, 0.05). 'ctype' is then used to specify what type of test we want to use. In this case, we are using the Bonferroni Multicomparison Test.

Additionally, we can substitute in:

'hsd' or 'tukey-kramer' for the Tukey's honest significant difference criterion.

'lsd' for Tukey's least significant difference procedure.

'dun-sidak' which uses critical values from the t-distribution after an adjustment for multiple comparisons proposed by Dunn and proved accurate by Sidak (hence, Dunn-Sidak).

'scheffe' which uses critical values from Scheffe's S-procedure, derived from the F-distribution.

OUTPUT:

'c' -- matrix of pairwise comparison results. Also contains an interactive graph of the estimates with comparison intervals around them (shown below).

'm' -- matrix in which the 1st column contains estimated values of **mean** for each group, and the 2nd column contains their respective st. errors.

'h' -- returns a handle to the comparison graph

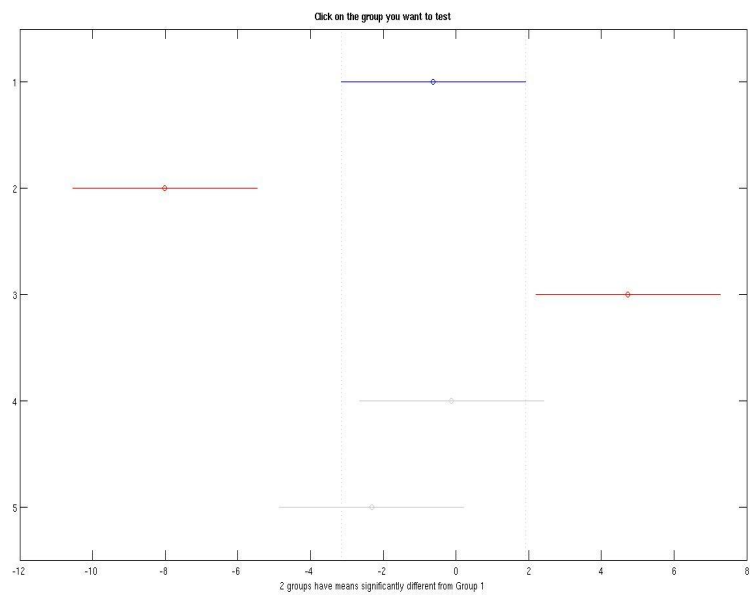
'nms' -- a cell array with one row for each group, containing the names of the groups.

For more information, please see [this Matlab Tutorial Link on Multicomparison Tests](#).

This ANOVA table is produced:

ANOVA Table					
Source	SS	df	MS	F	Prob>F
Columns	505.894	4	126.473	15.57	1.62141e-06
Error	203.104	25	8.124		
Total	708.998	29			

Along with this Multicompare graph:



Text version of MATLAB code

%Lines '5-8' are used to clear any stored data which may conflict with this
%program including overlapping variable names or not applicable data which
%could slow down run-time.

```
clc %Clear command screen
clf %Clear figures
clear all %Remove items from workspace, free up system memory
close all hidden %Close all windows from previous iteration of program
```

```
%List of drugs (below) which are formatted with the name of the drug/drug
%option on the left followed by the dependent variable or results of the
%study on the right. Data entered this way are stored as 'n-by-1' vectors.
placebo = [-1.12 -.98 .23 -1.21 -1.11 .45];
DrugOptA = [-6.02 -8.78 -7.48 -9.40 -8.35 -8.02];
DrugOptB = [5.70 4.86 5.95 3.79 3.57 4.54];
DrugOptC = [5.07 2.91 -9.48 -0.51 3.66 -2.36];
DrugOptD = [1.54 -0.79 -5.60 -2.53 -0.48 -6.07];
```

```
%Combine the vectors above into a single array called 'data' where each row
%contains the result of a single drug or drug option
data(:,1) = placebo;
data(:,2) = DrugOptA';
data(:,3) = DrugOptB';
data(:,4) = DrugOptC';
data(:,5) = DrugOptD';
```

```
[p, table, stats] = anova1(data);
[c, m, h, nms] = multcompare(stats,'alpha',.05,'ctype','bonferroni');
```