

How Prism 6 computes multiple comparisons tests following ANOVA (one-way and two-way)

Prism 6 can perform many kinds of multiple comparisons testing. All are standard, so there should be no surprises in this document, which reviews exactly how Prism does the calculations.

This document was created January 2011. Updates will be posted to:

<http://www.graphpad.com/faq/viewfaq.cfm?faq=1688>

How the family is defined?

One-way ANOVA

Number of families = 1

Compare each column mean with every other column mean

Number of comparisons within family = $N * (N - 1)/2$, where N is the number of data set columns

Compare each column mean with the control column mean

Number of comparisons within family = $(N - 1)$, where N is the number of data set columns.

Compare selected column means

User selects precisely which comparisons to make, so specifies the number of comparisons.

Two-way: Compare column means (main column effect)

Number of families = 1

Compare each column mean with every other column mean

Number of comparisons within family = $N * (N - 1)/2$, where N is the number of levels of column factor.

Compare each column mean with the control column mean

Number of comparisons within family = $(N - 1)$, where N is the number of levels of column factor.

Two-way: Compare row means (main row effect)

Number of families = 1

Compare each row mean with every other row mean

Number of comparisons within family = $N * (N - 1)/2$, where N is the number of levels of row factor.

Compare each row mean with the control row mean

Number of comparisons within family = $(N - 1)$, where N is the number of levels of row factor.

Two-way: Within each row, compare columns (simple effect within row)

The user can choose to define all the comparisons to be one family, or to create one family per row.

One family for all comparisons (conservative)

Number of families = 1

Compare each cell mean with every other cell mean on that row

Number of comparisons within family = $M * (N * (N - 1) / 2)$, where M is the number of levels of row factor and N is the number of levels of column factor.

Compare each cell mean with the control cell mean on that row

Number of comparisons within each family = $M * (N - 1)$, where M is the number of levels of row factor and N is the number of levels of column factor.

One family per row (recommended)

Number of families = number of levels of row factor.

Compare each cell mean with every other cell mean on that row

Number of comparisons within each family = $N * (N - 1) / 2$, where N is the number of levels of column factor.

Compare each cell mean with the control cell mean on that row

Number of comparisons within each family = $(N - 1)$, where N is the number of levels of column factor.

Two-way: Within each column, compare rows (simple effect within column)

The user can choose to define all the comparisons to be one family, or to create one family per column.

One family for all comparisons (conservative)

Number of families = 1

Compare each cell mean with every other cell mean on that column

Number of comparisons within family = $M * (N * (N - 1) / 2)$, where M is the number of levels of column factor and N is the number of levels of row factor.

Compare each cell mean with the control cell mean on that column

Number of comparisons within each family = $M * (N - 1)$, where M is the number of levels of column factor and N is the number of levels of row factor.

One family per column (recommended)

Number of families = number of levels of column factor.

Compare each cell mean with every other cell mean on that column

Number of comparisons within each family = $N * (N - 1)/2$, where N is the number of levels of row factor.

Compare each cell mean with the control cell mean on that column

Number of comparisons within each family = $(N - 1)$, where N is the number of levels of row factor.

Two-way: Compare each cell mean with the other cell mean in that row

Number of families = 1

Number of comparisons within family = the number of levels of row factor.

Two-way: Compare each cell mean with the other cell mean in that column

Number of families = 1

Number of comparisons within family = the number of levels of column factor.

Two-way: Compare all cell means regardless of rows and columns

Number of families = 1

Compare each cell mean with every other cell mean

Number of comparisons within family = $N * (N - 1)/2$, where N is the number of levels of row factor multiplied by the number of levels of column factor.

Compare each cell mean with the control cell mean

Number of comparisons within family = $(N - 1)$, where N is the number of levels of row factor multiplied by the number of levels of column factor.

How the appropriate error term is defined (MS(Error), DF(Error))?

One-way ANOVA (not repeated measures)

$$MS(\text{Error}) = MS(\text{Residual})$$

$$DF(\text{Error}) = DF(\text{Residual})$$

One-way ANOVA (repeated measures)

$$MS(\text{Error}) = MS(\text{Residual})$$

$$DF(\text{Error}) = DF(\text{Residual})$$

Yes, it is the same equation as for ordinary (not repeated measures) ANOVA, but the MS(residual) is computed differently with repeated measures.

Two-way ANOVA (not repeated measures; both factors are between-subjects)

For all kind of comparisons (main effects, simple effects) the MS_{Error} and DF_{Error} are defined as follows:

$$MS(\text{Error}) = MS(\text{Residual})$$

$$DF(\text{Error}) = DF(\text{Residual})$$

Two-way repeated measures ANOVA (both factors are within-subjects)

Compare column means (main column effect)

$$MS(\text{Error}) = MS(\text{Column Factor} \times \text{Subjects})$$

$$DF(\text{Error}) = DF(\text{Column Factor} \times \text{Subjects})$$

Compare row means (main row effect)

$$MS(\text{Error}) = MS(\text{Row Factor} \times \text{Subjects})$$

$$DF(\text{Error}) = DF(\text{Row Factor} \times \text{Subjects})$$

Compare cell means

$$MS(\text{Error}) = MS(\text{Column Factor} \times \text{Row Factor} \times \text{Subjects}) = MS(\text{Residuals})$$

$$DF(\text{Error}) = DF(\text{Column Factor} \times \text{Row Factor} \times \text{Subjects}) = DF(\text{Residuals})$$

Two-way repeated measures ANOVA. Column factor is between-subjects, Row factor is within-subjects (repeated measures)

Compare column means (main column effect), within each row, compare columns (simple effect within row), compare each cell mean with the other cell mean in that row

$$MS(\text{Error}) = (SS(\text{Subjects}) + SS(\text{Residuals})) / (DF(\text{Subjects}) + DF(\text{Residuals}))$$

$$DF(\text{Error}) = (DF(\text{Subjects}) + DF(\text{Residuals}))$$

Compare row means (main row effect), within each row, compare columns (simple effect within row), compare each cell mean with the other cell mean in that column

$$MS(\text{Error}) = MS(\text{Residuals})$$

$$DF(\text{Error}) = DF(\text{Residuals})$$

Compare all cell means regardless of rows and columns

$$MS(\text{Error}) = (SS(\text{Subjects}) + SS(\text{Residuals})) / (DF(\text{Subjects}) + DF(\text{Residuals}))$$

$$DF(\text{Error}) = (DF(\text{Subjects}) + DF(\text{Residuals}))$$

Multiple comparisons tests

Definitions:

Alpha = Significance level, set by user, often to 0.05.

K = Number of pairwise comparisons within family (Defined above)

M = Number of different means within each family

MS(Error) = Appropriate mean square error. Defined above.

DF(Error) = Number of degrees of freedom used in computing the MS(error), defined above.

N1 and N2 are calculated as follows:

Comparing column/row means

N1 = number of all values within column/row 1

N2 = number of all values within column/row 2

Compare cell means

N1 = number of values within cell 1

N2 = number of values within cell 2

Statistical functions: PFromQTukey, QFromPTukey, PFromQDunnett, QFromPDunnett, PFromT, and TFromP convert from statistical ratios to probabilities and from probabilities to ratios.

Adjusted vs. not adjusted P values, refers to adjustment for multiple comparisons. See:

<http://www.graphpad.com/faq/viewfaq.cfm?faq=1518>

Correct for multiple comparisons. Compute CI and significance

Bonferroni test

$$SEDifference = \sqrt{MS(Error) * \left(\frac{1}{N_1} + \frac{1}{N_2}\right)}$$

$$t = \frac{|Mean_1 - Mean_2|}{SEDifference}$$

$$PValue(adjusted) = PFromT(t, DF(Error)) * K$$

$$CIDifference = |Mean_1 - Mean_2| \pm TFromP(Alpha/K, DF(Error)) * SEDifference$$

Tukey test

$$SEDifference = \frac{1}{\sqrt{2}} \sqrt{MSerror * \left(\frac{1}{N_1} + \frac{1}{N_2}\right)}$$

$$q = \frac{|Mean_1 - Mean_2|}{SEDifference}$$

$$CIDifference = |Mean_1 - Mean_2| \pm QFromPTukey(Alpha, DF, M) * SEDifference$$

$$PValue(adjusted) = PFromQTukey(q, DF, M)$$

Dunnet test

$$SEDifference = \sqrt{MSerror * \left(\frac{1}{N_1} + \frac{1}{N_2}\right)}$$

$$q = \frac{|Mean_1 - Mean_2|}{SEDifference}$$

$$CIDifference = |Mean_1 - Mean_2| \pm QFromPDunnet(Alpha, DF, M) * SEDifference$$

$$PValue(adjusted) = PFromQDunnet(q, DF, M)$$

Note that for historical reasons, the standard error of the difference is defined differently for the Dunnett and Tukey tests. The Dunnett test definition lacks the square root of 2 term.

$$q = \frac{|Mean_1 - Mean_2|}{SEDifference}$$

$$CIDifference = |Mean_1 - Mean_2| \pm QFromPDunnett(Alpha, DF, M) * SEDifference$$

$$PValue(adjusted) = PFromQDunnett(q, DF, M)$$

Correct for multiple comparisons. Compute significance only

Newman-Keuls test

Let's assume that family consists of 5 samples with the following means: M1 ... M5. Let's assume that after ranking (sorting in decreasing order) means are ordered as follows: M2, M5, M1, M4, M3.

Then the pairs are considered in the following order: largest vs. smallest, largest vs. next smallest ... largest vs. next largest, next largest vs. smallest ... next smallest vs. smallest.

In our example the order of pairs being considered looks like: (M3 vs. M2), (M3 vs. M5), (M3 vs. M1), (M3 vs. M4), (M4 vs. M2), (M4 vs. M5), (M4 vs. M1), (M1 vs. M2), (M1 vs. M5), (M5 vs. M2).

For each pair let's designate B as the number of Means within this range (e.g. for second pair (M3, M5) P = 4).

After that the pairwise comparisons are calculated as follows (in the order defined above):

$$SEDifference = \frac{1}{\sqrt{2}} \sqrt{MSerror * \left(\frac{1}{N_1} + \frac{1}{N_2}\right)}$$

$$q = \frac{|Mean_1 - Mean_2|}{SEDifference}$$

$$PValue(adjusted) = PFromQTukey(q, DF, B)$$

The calculations are the same as Tukey test, with the one exception. The Tukey test is based on M (number of all means within family), while the Newman-Keuls test is based on B (the number of means within this range).

Another important rule is that if no difference is found between two means then it's concluded that no difference exists between any means enclosed by those two, and such differences are not tested for.

Holm-Sidak test

First for each pair within family unpaired t test is calculated as follows:

$$SEDifference = \sqrt{MS(Error) * \left(\frac{1}{N_1} + \frac{1}{N_2}\right)}$$

$$t = \frac{|Mean_1 - Mean_2|}{SEDifference}$$

$$PValue = PFromT(t, DF(Error))$$

Then p values (two-tailed) are sorted in increasing order and statistical significance is calculated by the following algorithm.

- 1) Start with the smallest P value. Set K equal to the number of comparisons.
- 2) Compare the P value to this value: $(1 - \text{Alpha})^{\frac{1}{K}}$
- 3) If the P value is less than or equal to the computed value, that difference is considered to be statistically significant. Go back to step 2, decrease the value of K by 1, and test the next largest P value .
- 4) If the P value is greater than or equal to the computed value, then that difference is considered to not be statistically significant. Don't test any more P values. All the rest are considered to be not statistically significant

Fisher's LSD (Least Significant Difference) test

$$SEDifference = \sqrt{MS(Error) * \left(\frac{1}{N_1} + \frac{1}{N_2}\right)}$$

$$t = \frac{|Mean_1 - Mean_2|}{SEDifference}$$

$$PValue(NotAdjusted) = PFromT(t, DF(Error))$$

$$CIDifference = |Mean_1 - Mean_2| \pm TFromP(\text{Alpha}, DF(Error)) * SEDifference$$

Prism does the *unprotected* Fisher's LSD test, which means the individual comparisons are made without regard to the overall ANOVA results. In contrast, the *protected* test only does individual comparisons if the overall P value is less than a preset threshold.

Read more

Difference between q and t ratios: <http://www.graphpad.com/faq/viewfaq.cfm?faq=540>

How to compute the critical values of q for Tukey and Dunnett test:

<http://www.graphpad.com/faq/viewfaq.cfm?faq=1517>

How to compute critical values of the t distribution: <http://www.graphpad.com/faq/viewfaq.cfm?faq=1431>

The post test for trend: <http://www.graphpad.com/faq/viewfaq.cfm?faq=1544>

Multiple comparisons after repeated measures ANOVA, with or without assuming sphericity:

<http://www.graphpad.com/faq/viewfaq.cfm?faq=1609>

Dunn's multiple comparison test following the nonparametric Kruskal-Wallis ANOVA:

<http://www.graphpad.com/faq/viewfaq.cfm?faq=1156>

All the methods are explained in SE Maxwell and HD Delaney, *Designing Experiments and Analyzing Data*, Second edition, Lawrence Erlbaum, ISBN: 0-8058-3718-3