

Graphpad 分析教程-多因素方差分析

今天给大家介绍强大的 graphpad prism 8.0 是如何进行多因素方差分析的，我们以双（多）分析中的析因分析为例来介绍。

首先我们要明确，析因分析所关心的问题主要有两个：

- 两个或两个以上处理因素的各处理水平间均数有无差异？及主效应有无统计学意义
- 两个或两个以上处理因素之间有无交互作用

话不多说，我们直接开始：

1. 为了方便仍然以 sample data 来介绍，如下图。当然也可以按照第二种选择，设定平行样本数，即可得到数据表。

The image displays two screenshots of the Graphpad Prism 8.0 software interface, specifically the 'New table & graph' dialog box. Both screenshots show the 'Grouped' option selected in the 'Data table' section. The left screenshot highlights 'Start with sample data to follow a tutorial' in the 'Data table' section. The right screenshot highlights 'Enter and plot a single Y value for each point' in the 'Options' section, and 'replicate values in side-by-side subcolumns' in the 'Data table' section, with a red box and the text '平行样本数为5' next to it. The dialog box also includes a 'Table format' table and a 'Learn more' link.

Table format	A		B	
Grouped	Control	Treated	Control	Treated
1	A:Y1	A:Y2	B:Y1	B:Y2
2	Male			
3	Female			

Table format: Grouped		Group A					Group B					
		Wild-type cells					GPP5 cell line					
		A:1	A:2	A:3	A:4	A:5	B:1	B:2	B:3	B:4	B:5	C:1
1	Serum starved	34	36	41	40	43	98	87	95	99	88	
2	Normal culture	23	19	26	29	25	32	29	26	33	30	
3	Title											
4	Title											
5	Title											
6	Title											
7	Title											
8	Title											
9	Title											
10	Title											
11	Title											
12	Title											
13	Title											
14	Title											
15	Title											
16	Title											
17	Title											
18	Title											
19	Title											
20	Title											
21	Title											
22	Title											
23	Title											
24	Title											

各组设置的不同分析因素

各组设置的不同分析因素下的相对应平行样本数据

How the data are organized
 The columns represent two cell lines. The rows represent two treatments. Within each treatment for each cell line, five replicate values are entered into subcolumns. This experiment has no matching or repeated measures. For this reason, it is OK that one of the values is missing (so its spot is simply left blank).

Note that, unlike other programs, Prism does not use grouping variables. Instead the treatments are defined by rows, and the groups by columns.

The goals

- To assess whether the difference between cell lines is more than expected by chance.
- To assess whether the difference between treatments is more than expected by chance.
- To assess whether the difference between treatments is consistent for each cell line.
- To compute a 95% confidence interval for the difference between cell lines, for each treatment.

How to perform two-way ANOVA
 Click Analyze, choose "Two-way ANOVA" from the list of Grouped analyses, and accept all the default choices on the dialog. Click the link below for detailed instructions, and to learn about two-way ANOVA.

[Step by step instructions for performing two-way ANOVA](#)

2. 点击 analyze 或左侧 results 的 new analysis 后进入 creat a new analysis, 按下图示选择 grouped Analyses 下的 two-way ANOVA(or mixed model) , 注意不要漏选 data sets 里的 A 和 B , 得到新对话框

Data to analyze

Table: Two-way ANOVA, not RM

Type of analysis

Which analysis?

- Transform, Normalize...
- XY analyses
- Column analyses
- Grouped analyses**
 - Two-way ANOVA (or mixed model)**
 - Three-way ANOVA (or mixed model)
 - Row means with SD or SEM
 - Multiple t tests - one per row
- Contingency table analyses
- Survival analyses
- Parts of whole analyses
- Multiple variable analyses
- Nested analyses
- Generate curve
- Simulate data
- Recently used

Analyze which data sets?

- A: Wild-type cells
- B: GPP5 cell line

➔

RM Design | RM Analysis | Factor names | Multiple Comparisons | Options | Residuals

Data arrangement

Table format: Grouped	Group A		Group B		Group C	
	Title	Title	Title	Title	Title	Title
1	A:Y1	A:Y2	B:Y1	B:Y2	C:Y1	C:Y2
2						
3						
4						

Matching by which factor(s)?

- Each column represents a different time point, so matched values are spread across a row.
- Each row represents a different time point, so matched values are stacked into a subcolumn.

Assume sphericity (equal variability of differences)?

- No. Use the Geisser-Greenhouse correction. Recommended.
- Yes. No correction.

Based on your choices (on all tabs), Prism will perform:

- Ordinary two-way ANOVA

2.1 选择 RM design : 注意勾选不同引起的变化 (主要下图中的绿色框 , 注意行列、纵列、组列) 。

RM Design | RM Analysis | Factor names | Multiple Comparisons | Options | Residuals

Data arrangement

Table format:		Group A		Group B		Group C	
Grouped		Title		Title		Title	
		A:Y1	A:Y2	B:Y1	B:Y2	C:Y1	C:Y2
1	Title						
2	Title						
3	Title						
4	Title						

Matching by which factor(s)? 单选或复选

Each column represents a different time point, so matched values are spread across a row.

Each row represents a different time point, so matched values are stacked into a subcolumn.

Assume sphericity (equal variability of differences)?

No. Use the Geisser-Greenhouse correction. Recommended.

Yes. No correction.

Based on your choices (on all tabs), Prism will perform:
- Ordinary two-way ANOVA

RM Design | RM Analysis | Factor names | Multiple Comparisons | Options | Residuals

Data arrangement

Table format:		Group A		Group B		Group C	
Grouped		Time1		Time2		Time3	
		A:Y1	A:Y2	B:Y1	B:Y2	C:Y1	C:Y2
1	Title						
2	Title						
3	Title						
4	Title						

Matching by which factor(s)?

Each column represents a different time point, so matched values are spread across a row.

Each row represents a different time point, so matched values are stacked into a subcolumn.

Assume sphericity (equal variability of differences)?

No. Use the Geisser-Greenhouse correction. Recommended.

Yes. No correction.

Based on your choices (on all tabs), Prism will perform:
- RM two-way ANOVA, matched values are spread across a row.
- Sidak's multiple comparisons test, with a single pooled variance.

Because your repeated measures factor has only two levels, the concept of sphericity doesn't apply.

RM Design | RM Analysis | Factor names | Multiple Comparisons | Options | Residuals

Data arrangement

Table format:		Group A		Group B		Group C	
Grouped		Title		Title		Title	
		A:Y1	A:Y2	B:Y1	B:Y2	C:Y1	C:Y2
1	Time1						
2	Time2						
3	Time3						
4	Time4						

Matching by which factor(s)?

Each column represents a different time point, so matched values are spread across a row.

Each row represents a different time point, so matched values are stacked into a subcolumn.

Assume sphericity (equal variability of differences)?

No. Use the Geisser-Greenhouse correction. Recommended.

Yes. No correction.

Based on your choices (on all tabs), Prism will perform:
- RM two-way ANOVA, matched values are stacked into a subcolumn.
- Sidak's multiple comparisons test, with a single pooled variance.

Because your repeated measures factor has only two levels, the concept of sphericity doesn't apply.

RM Design | RM Analysis | Factor names | Multiple Comparisons | Options | Residuals

Data arrangement

Table format:		Group A		Group B		Group C	
Grouped		Title		Title		Title	
		A:Y1	A:Y2	B:Y1	B:Y2	C:Y1	C:Y2
1	Title						
2	Title						
3	Title						
4	Title						

Matching by which factor(s)?

Each column represents a different time point, so matched values are spread across a row.

Each row represents a different time point, so matched values are stacked into a subcolumn.

Assume sphericity (equal variability of differences)?

No. Use the Geisser-Greenhouse correction. Recommended.

Yes. No correction.

Based on your choices (on all tabs), Prism will perform:
- RM two-way ANOVA with the Geisser-Greenhouse correction, matched values are both stacked and spread across a row.
- Sidak's multiple comparisons test, with individual variances computed for each comparison.

2.2 进入 RM analysis 菜单，默认情况下选择如图，一般选系统推荐的，但注意有的时候推荐选项不是最常用选项，按下图选择

RM Design **RM Analysis** Factor names Multiple Comparisons Options Residuals

Analyses of repeated measures data can be reported in two ways.

- ANOVA (partition sum-of-squares). This is the same as the general linear model (GLM).
- Mixed-effects model. This uses the restricted maximum likelihood method.

If there are no missing values, the two approaches give the same main results (F and P values). But the methods are very different, so the other reported results differ.

Analyze using which method

- Repeated measures ANOVA (based on GLM).
Same as Prism 7 and earlier.
Requires balanced data (no missing values).
- Mixed-effects model.
Results are presented in a format different than ANOVA.
Works fine with missing values.
- It depends.
Use ANOVA if there are no missing values.
Use mixed-effects model if there are missing values.

What to do if a random effect is zero (or negative)?

- Remove term(s) from model and fit a simpler model **recommended**.
- Fit the full model anyway (corresponds to NOBOUND parameter in SAS).

Make these choices the default for future ANOVAs (One-, Two- and Three-way).

2.3 进入 multiple comparisons，选择第三个「每行进行组别比较，也就是在每一行里的个组别间进行比较」。

第一个为「无多重比较，各组间各列差异」

第二个为「各组间行间差异」

第四个为「各组间行间列间的复合差异」（见下图）

RM Design | RM Analysis | Factor names | **Multiple Comparisons** | Options | Residuals

What kind of comparison?

Compare each cell mean with the other cell mean in that row < >

No multiple comparisons —

Compare each cell mean with the other cell mean in that column —

Compare each cell mean with the other cell mean in that row —

Compare cell means regardless of rows and columns 四

1	Mean	←	→	Mean
2	Mean	←	→	Mean
3	Mean	←	→	Mean

如何比较

How many comparisons?

Compare each column mean with every other column mean. 无对照两两比较

Compare each column mean with the control column mean. 设置对照后比较

Control column: Group A : Wild-type cells

Which test?

Use choices on the Options tab to choose the test, and to set the defaults for future ANOVAs.

2.4 然后选择 options，按图示选择推荐选项，切记拿不准的选项就默认。

RM Design | RM Analysis | Factor names | Multiple Comparisons | **Options** | Residuals

Multiple comparisons test

Correct for multiple comparisons using statistical hypothesis testing. Recommended.
 Test: Sidak (more power, recommended)

Correct for multiple comparisons by controlling the False Discovery Rate.
 Test: Two-stage step-up method of Benjamini, Krieger and Yekutieli (recommended)

Don't correct for multiple comparisons. Each comparison stands alone.
 Test: Fisher's LSD test

Multiple comparisons options

Swap direction of comparisons (A-B) vs. (B-A).

Report multiplicity adjusted P value for each comparison.
 Each P value is adjusted to account for multiple comparisons.

Family-wise significance and confidence level: 0.05 (95% confidence interval)

Graphing options

Graph confidence intervals.

Additional results

Narrative results.

Show cell/row/column/grand means.

Report goodness of fit.

Output

Show this many significant digits (for everything except P values): 4

P value style: GP: 0.1234 (ns), 0.0332 (*), 0.0021 (**), 0.0001 (***) N = 6

Make options on this tab be the default for future Two-Way ANOVAs.

3. 最后分析结果（看每种情况下不同的 P 值）：

ANOVA results | Multiple comparisons

Source of Variation	% of total variation	P value	P value summary	Significant?	Geisser-Greenhouse's epsilon
Row Factor	49.25	<0.0001	****	Yes	1.000
Column Factor	29.49	0.0001	***	Yes	1.000
Interaction: Row Factor x Column Factor	19.54	0.0001	***	Yes	1.000
Interaction: Row Factor x Subject	0.08267				
Interaction: Column Factor x Subject	0.5123				
Subject	0.7375				

ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
Row Factor	7566	1	7566	F (1, 4) = 2383	P<0.0001
Column Factor	4530	1	4530	F (1, 4) = 230.2	P=0.0001
Interaction: Row Factor x Column Factor	3001	1	3001	F (1, 4) = 198.4	P=0.0001
Interaction: Row Factor x Subject	12.70	4	3.175		
Interaction: Column Factor x Subject	78.70	4	19.68		

各变量对总体方差变异的影响

p < 0.05

$P < 0.05$ 说明变量数据中存在具有统计学意义的显著性差异，如果需要明确哪

一对或者几对数据具有差异，则需要查看 Multiple comparisons 里面的结果。

2way ANOVA		Multiple comparisons					
Compare each cell mean with the other cell mean in that row							
Number of families	1						
Number of comparisons per family	2						
Alpha	0.05						
Sidak's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Significant?	Summary	Adjusted P Value		
Wild-type cells - GPP5 cell line							
Serum starved	-54.60	-60.95 to -48.25	Yes	****	<0.0001 < 0.05		
Normal culture	-5.600	-11.95 to 0.7510	No	ns	0.0879 > 0.05		
Test details	Mean 1	Mean 2	Mean Diff.	SE of diff.	N1	N2	t
Wild-type cells - GPP5 cell line							
Serum starved	38.80	93.40	-54.60	2.575	5	5	21.20
Normal culture	24.40	30.00	-5.600	2.575	5	5	2.175

4. 绘图：可选用散点图（数据较少）或误差线柱状图（数据较多）

