

## **Graph**Pad



Analyze, graph and present your scientific work.

## GraphPad Prism 9 常用数据统计分析

### 科研技能 单元课 03 增刊

■ 针对特定科研技能的专项突破

上海尤里卡信息科技有限公司 版权所有





Location of transferrin receptors





## 科研技能 单元课

www.helixlife.cn/





解螺旋 | 陪伴医生科研成长

## 科研技能 单元课

数据统计分析概述

在对于统计学方面而言,可谓是繁琐而冗长。所以在这里仅对一些常用的统计方法及软件应用进行简单教学,下面的概述更偏向于应用:

▶ 参数检验与非参数检验:

1、参数检验是针对参数做的假设

非参数检验是针对总体分布情况做的假设,这是区分的一个重要特征

- 2、参数检验要利用到总体的信息(总体的分布、总体的一些参数特征,如方差),以总体分布和样本信息对总体参数做出推断 非参数检验不需要利用总体信息,以样本信息对总体分布做出推断
- ▶ 数据表示方法:
- ① 计量资料符合正态分布采用均数±标准差表示,例:本班男生身高为175±3cm
- ② 资料不符合正态分布采用中位数(四分位数间距)表示,例:该组患者胆红素水平为50(21,70) µM
- ③ 计数资料采用率或构成比表示,例: 该组患者复发率为30%
- > 假设检验: (假设检验的选择应该根据数据类型进行)
- ① 符合正态分布与方差齐性的两组间计量资料比较采用两独立样本t检验
- ② 不符合正态分布与方差齐性的两组计量资料比较采用非参数Mann-Whitney U检验
- ③ 配对设计计量资料比较,差值符合正态分布采用配对t检验;差值不符合正态分布采用非参数Willcoxon秩和检验(Sign检验)
- ④ 多组间计量资料比较符合条件采用单因素设计方差分析;组间两两比较采用LSD法(SNK法/Bonferroni法/Duncan法);多组间计量资料比较不符合应用条件采用Kruskal-Wallis H法
- ⑤ 配伍组设计计量资料比较符合条件采用随机区组设计方差分析;不符合条件采用非参数Friedman检验
- ⑥ 等级资料组间比较采用非参数秩和检验
- ⑦ 成组四格表计数资料符合条件采用Pearson卡方检验;不符合条件采用Pearson连续校正卡方检验或Fisher 's确切概率法
- ⑧ 成组设计R\*C表符合条件采用Pearson卡方检验;不符合条件采用MonteCarlo近似确切概率法

## 科研技能 单元课

数据统计分析概述

- 也可以根据下表对适用的统计方法进行选择
- 本课程仅对红字内容进行介绍



## 科研技能 单元课

数据统计分析概述

#### ▶ 相关分析

- ① 双变量正态分布资料采用Pearson相关系数
- ② 双变量非正态分布或等级资料采用Spearman相关系数
- ③ 当面临两组多变量数据,并希望研究两组变量之间的关系时,就要用到典型相关分析( canonical correlation analysis)

▶ 回归分析:

- ① 线性回归:因变量为计量资料的影响因素分析采用多元线性回归
- ② Logistic回归:二分类因变量影响因素分析采用二元Logistic回归(Binary Logistic)
- ③ 有序因变量影响因素分析采用有序逻辑回归(Ordinal Logistic)

④ 无序多分类计数资料因变量影响因素分析采用多项Logistic回归(Multinominal Logistic)

*▶ P*值:

P<0.05或P<0.01具有统计学意义;如有可能,请将P值报告为一个精确到有几位数的数值,而非一个不等式。例如说"P值为0.0234",而非"P<0.05"

## 科研技能 单元课

数据统计分析概述



Analyze Data  $\times$ ① GraphPad Prism 9里,在数据或者作图页面,点击顶 (2) Use: Built-in analysis  $\sim$ Recently used 最近使用 (1)Which analysis? Analyze which data sets? 部菜单栏的Analyze按钮,即可调出数据分析功能窗口 Search.  $\sim$ Table: Data 1 Transform, Normalized 转化,标准化 (2)(2) Recently used A:si-NC (3) XY analyses XY分析 根据需要的分析功能对数据进行选择 (2) H Transform, Normalize... B:si-BCRT1 XY analyses Column analyses (4)Column analyses Column分析 H Grouped analyses 选择需要分析的数据 (3) Contingency table analyses (5) Grouped analyses 成组分析 Survival analyses Parts of whole analyses Hultiple variable analyses 列联表分析 Contingency table analyses (6)(4) 选择OK Nested analyses Generate curve Survival analyses 生存分析 (7)🗄 Simulate data 部分整体分析 (8)Parts of whole analyses Multiple variable analyses 多变量分析 (9)Nested analyses 嵌套分析 (10)曲线生成 (11) Generate curve 模拟数据 Deselect All (12) Simulate data 4

Cancel

Help

ОК

## 科研技能 单元课

www.helixlife.cn/

### 数据统计分析概述

- GraphPad Prism的数据分析功能强大,且针对不同的分析,还自带示例供参照学习:
- ①在新建数据表格的窗口中,可选择某一分析的示例数据



#### ② 点击"Create"后,新创建数据页面中即给出示例数据和说明



05

## 科研技能 单元课

www.helixlife.cn/

【课程目录



## 科研技能 单元课

### 假设检验

- ▶ 假设检验(hypothesis test)主要是验证所选的模型和所解释的公式,在结构上、形式上、变化方向上是否能代表客观情况。一般采用最小平方法解模时必须 进行统计假设检验,即应用统计推断的假设检验原理,通过特定的统计量进行各种显著性检验。假设检验的本质是利用样本对假设的真假进行判断
- ▶ 要进行统计假设的检验,必须利用各种不同的判据,即利用规则来选择。先假设总体某项假设成立,计算其会导致什么结果产生。若导致不合理现象产生,则拒绝原先的假设。若并不导致不合理的现象产生,则不能拒绝原先假设,从而接受原先假设。假定样本数据来自具有特定分布的总体,则使用参数检验。如果不能对数据集作出必要的假设,则使用非参数检验

▶ 假设检验的一般归类:

① 如果已知分布的均值和方差,则可以构造标准正态分布,即u检验

② 如果已知分布的均值,不知道方差,则可以用方差的无偏估计代替方差,即t检验法

③ 如果均值、方差都不知,利用样本的方差,构造卡方分布,即卡方检验

④ 如果均值、方差都不知,存在两个分布的比较,构造F分布,查F分布表,即F检验

# 科研技能单元课

正态检测

### 1 在选

#### 在选择分析方法前,需要对数据的分布及方差进行检测;点击Analyze按钮,调出数据分析窗口

Analyze Data	×	Parameters: Normality and Lognormality Tests $\qquad \qquad \qquad$	
Use: Built-in analysis	~		① 远拴止念位验
		Which distribution(s) to test?	
Which analysis? Ana	alyze which data sets?	Normal (Gaussian) distribution	② 选择要检验的数据
Search Tab	ble: Data 4 🗸 🗸	3 Lognormal distribution	
		Compute the relative likelihood of sampling from a Gaussian (normal) vs.	③ 选择要检验什么分布
		a lognormal distribution (assuming no other possibilities)	
Transform, Normalize	<u>_B:B</u>	Methods to test distribution(s)	
Transform concentrations (V)		Anderson-Darling test	④ 选择检验的方法
Normalize		✓ D'Agostino-Pearson omnibus normality test	
Prune rows		(4) √ Shapiro-Wilk normality test	4.77月—— 种数据 网种方法的计算结果
Remove baseline and column math		Kolmogorov-Smirnov normality test with Dallal-Wilkinson-Lilliefor P value	
Transpose X and Y			
Fraction of total		Graphing options	大同小异: 虽然GraphPad Prism不推荐
XY analyses		✓ Create a QQ plot	
Column analyses		Subcolumns	
t tests (and nonparametric tests)		Average the replicates in each row, and then perform the calculation	用Kolmogorov-Smirnov法,小样本多
One-way ANOVA (and honparametric or		for each column	
Descriptive statistics		Perform calculations on each subcolumn separately	
Normality and Lognormality Tests		Treat all the values in all subcolumns as single set of data	采用Snapiro-wilk normality test
Frequency distribution		Calculations	
ROC Curve		Significance level (alpha) 0.05	
Bland-Altman method comparison			
Identify outliers		output	
		Show this many significant digits (for everything except P values):	
Two-way ANOVA (or mixed model)		P value style: GP: 0.1234 (ns), 0.0332 (*), 0.0021 (**), V N = 6	
	Soloct All Decelect All		
	Deselect All	Make these choices the default for future analyses.	
	Help Cancel OK	Learn Cancel OK	

## 科研技能 单元课

#### 正态检测

✓ Results	>
Mann-Whitney test of Data 4	
Normality and Lognormality Tests of Data 4	

① New Analysis...

在左边导航栏里找到结果列表:在上方Results中第二栏为检测结果,单击结果如右图 (1)

>>

四种方法的检验结果 (2)

✓ 常用的是 Shapiro-Wilk 检验(SW)和Kolmogorov-Smirnov 检验(KS)

✓ 大多数情况下,三种方法的结论一致;不一致时,一般小样本(n<30,有争议,也

有人认为n<200,还有人认为是n<2000)以SW法计算结果为准;大样本以KS法计 算结果为准

- ③ 一般认为,当*p* value > 0.1 (也有人认为*p* > 0.05) 时,数据服从正态分布
  - ✓ 这里的0.1 (或0.05) 是置信度,理解: 认为10% (5%) 为小概率事件,即比这个 几率小的事件,我们认为在实验中不会发生。设定p > 0.1认为数据符合正态分布, 对正态分布的检验更保守

81	Normality and Lognormality Tests	Α	В
	Tabular results	А	В
1	Test for normal distribution		
2	Anderson-Darling test		
3	A2*	0.5845	0.1982
4	P value	0.1225	0.8649
5	Passed normality test (alpha=0.05)'	Yes	Yes
6	P value summary	ns	ns
7			
8	D'Agostino & Pearson test		
9	К2	6.224	0.6972
10	P value	0.0445	0.7057
11	Passed normality test (alpha=0.05)'	No	Yes
12	P value summary	*	ns
13			
14	Shapiro-Wilk test		
15	W	0.9653	0.9747
16	P value	0.0857	0.8938
17	Passed normality test (alpha=0.05)'	Yes	Yes
18	P value summary	ns	ns
19			
20	Kolmogorov-Smirnov test		
21	KS distance	0.1025	0.1107
22	P value	>0.1000	>0.1000
23	Passed normality test (alpha=0.05)'	Yes	Yes
24	P value summary	ns	ns

# 科研技能单元课

### U检验

• 示例: 文献: Partner R, Jones B, Tee J, Francis P. Playing through the pain: The prevalence of perceived shoulder dysfunction in uninjured rugby players using the Rugby Shoulder Score. Phys Ther Sport. 2022 Jan 16;54:53-57. doi: 10.1016/j.ptsp.2022.01.001. Epub ahead of print. PMID: 35065316. Table 2.

#### 示例解读

- 多个两组数据
- 数据表现形式为均值士方差或中位数(四分位差)
- 每种分组依据不同,结果源于不同个体(非配对) , 两组数据在个数上可以不相等

	n	RSS Score
All Shoulders	139	35 ± 20
		23 (28)
No performance impact (Group 2)	60	40 (22)
Performance impact (Group 3)	17	61 (28)
		p = 0.02
Rugby League	125	25 (26)
Rugby Union	14	20 (25)
		p = 0.41
Professional	58	40 (39)
Amateur	81	20 (18)
		p = 0.02
Forwards	73	29 (29)
Backs	66	20 (21)
		p = 0.036
Previous Injury	48	48 (41)
No Previous injury	91	20 (13)
		p < 0.001

Table format:		Group A	Group B
Grouped		А	В
	×		
1	Title	42	81
2	Title	78	106
3	Title	14	27
4	Title	43	90
5	Title	62	115
6	Title	29	95
7	Title	68	70
8	Title	77	54
9	Title	38	25
10	Title	15	43
11	Title	11	89
12	Title	54	3
13	Title	59	31
14	Title	34	52
15	Title	57	71
16	Title	29	49
17	Title	51	80
18	Title	80	
19	Title	70	
20	Title	57	
21	Title	37	

#### 模拟数据:

No performance impact组的RSS

Score 组60个数据;

performance impact组17个数据

## 科研技能 单元课

### U检验

#### • 分析目标

- ✓ 数据总体上看,是对shoulders的多个一维分组,如:有或无功能影响;成员来自橄榄球联会或橄榄球联盟;专业运动员或业余运动员;前锋或后卫;受过 伤或没受过伤。所有的数据以均值±方差或中位数(四分位差)表示
- ✔ 从实验结果上看,需要每个分组是否存在显著性差异,比较一维两组独立样本,要看数据是否符合正态分布(本组数据组明显不符合),从而使用U检验



## 科研技能 单元课

www.helixlife.cn/

### U检验

• 数据录入及作图请查看单元课03《GraphPad Prism 9 作图教程》,本教程为单元课03增刊,旨在讲述在作图后如何进行统计分析,作图操作不再重复

1 Built in analysis

点击Analyze按钮,调出数据分析窗口

① 点击t test(在选择t检验时,也是点击此选项,

不要疑惑为什么做U检验时,点的是t检验)

② 勾选AB两组数据

③ 点击OK

Analyze Data					$\times$
Use:	Built-in analysis	$\sim$			
Which analysis?		Analyze which data	sets?		
Search		Table: Data 4		~	e
Recently used	^	A:A			
<ul> <li>□ Transform, Normalize Transform Transform concentrations ( Normalize Prune rows Remove baseline and colum Transpose X and Y Fraction of total</li> <li>■ XY analyses Column analyses Column analyses t tests (and nonparametric One-way ANOVA (and nonp One sample t and Wilcoxon Descriptive statistics Normality and Lognormality Frequency distribution ROC Curve Bland-Altman method comp Identify outliers Analyze a stack of P values Two-way ANOVA (or mixed The analyses Two-way ANOVA (or mixed The analyses The anal</li></ul>	x) nn math tests) 1 parametric or test Tests parison	B:B			
<	>	Select All		Deselect All	
		Help	Cancel	ОК	

## 科研技能 单元课

U检验

#### 2 参数设置1



Parameters: t tests (and Nonpara	Parameters: t tests (and Nonparametric Tests)		
Experimental Design Residuals Option	S		
What graphs to create?			
Residual plot	Homoscedasticity plot		
Correct model?	Equal variance?		
Residual ° ° ° ° °	Abs (Residual)		
QQ plot	Heatmap plot		
Normality?			
Diagnostics for residuals  Are the residuals Gaussian? Normality tests of Anderson-Darlin Kolmogorov-Smirnov.  Make options on this tab be the defa	g, D'Agostino, Shapiro-Wilk and ult for future tests.		
	Learn Cancel Ok	:	

参数设置包括实验设计,残差以及选项 ①实验设计 U检验选择非配对,非参数检验,M-W法 ②残差图选择(当前版本的Prism可以直接设置残差图) 勾选你需要的残差图种类

- (1) 实验设计
- (2) 非配对
- (3) 配对
- (4) 是否服从正态分布
- (5) 是,使用参数检验
- (6) 否,使用非参数检验
- (7) 选择检验方法
- (8) M-W法,检验两个总体的分布是否有显著的差异

(一般用这个)

(9) K-S法, 检验一个分析变量是否符合某种分布或

两个分析变量的分布是否相同

## 科研技能 单元课

U检验

#### 2 参数设置2

	Parameters: t tests (and Nonparametric Tests)	
	Experimental Design Residuals Options 3	
(1)	Calculations	
(1)	P value: One-tailed  Two-tailed (recommended)	
(2)	Report differences as: B - A	
(3)	Confidence level: 95% V	
	Definition of statistical significance: P < 0.05	
(4)	Graphing options	
	Graph differences (paired)	
	Graph ranks (nonparametric)	
	Graph correlation (paired)	
	✓ Graph CI of difference between medians	(0
(5)	Additional results	(9
(3)	Descriptive statistics for each data set	
	t test: Also compare models using AICc	
	Mann-Whitney: Also compute the CI of difference between medians	
	Assumes both distributions have the same shape.	(10
	Wilcoxon: When both values on a row are identical, use method of Pratt	
	If this option is unchecked, those rows are ignored.	
(c)	Output	
(6)	Show this many significant digits (for everything except P values):	
(7)	P value style: GP: 0.1234 (ns), 0.0332 (*), 0.0021 (**), I V N = 6	
(8)	Make options on this tab be the default for future tests.	
	Learn Cancel OK	)

Search ✓ ✓ Data Tables »
<ul> <li>Data Tables »</li> <li>Data 4</li> <li><i>New Data Table</i></li> <li>Info »</li> <li>Project info 1</li> </ul>
<ul> <li>Data 4</li> <li><i>New Data Table</i></li> <li>Info &gt;&gt;</li> <li>Project info 1</li> </ul>
<ul> <li>① New Data Table</li> <li>✓ Info</li></ul>
<ul> <li>Info »</li> <li>Project info 1</li> </ul>
<ol> <li>Project info 1</li> </ol>
<u> </u>
New Info
✓ Results »
Mann-Whitney test of Data 4
New Analysis
✓ Graphs »
📐 Homoscedasticity plot: Mann
New Graph
✓ Layouts »
New Layout

#### 参数设置包括实验设计,残差以及选项

①实验设计

U检验选择非配对,非参数检验,M-W法

②残差图选择(当前版本的Prism可以直接设置残差图)

勾选你需要的残差图种类

#### ③选项

选择结果展示的方式

④点击OK后,可出现分析结果以及上一步中勾选创建的图案

- (1) 选择单侧检验或是双侧检验, 默认双侧, 一般不建议修改
- (2) 选择比较的方式,谁与谁比
- (3)置信水平,一般是95%,即定义p<5%时认为有显著差异
- (4) 绘图选项,可选择绘制统计分析的图,可以不选
- (5) 补充结果,可以不选
- (6) 设置统计结果的有效数字
- (7) p值显示的方式
- (8) 勾选这个可将目前的选择设置为默认选项
- (9) 分析的结果
- (10) 残差图

## 科研技能 单元课

www.helixlife.cn/

#### U检验

#### 2 结果显示

Search	$\sim$	1	2 Mann-Whitney test	
✓ Data Tables	>>	(		
Data 4			[	
		1	Table Analyzed	Data 4
New Data Table	_	2		
v Info	»	3	Column B	В
(i) Project info 1		4	VS.	VS.
New Info		5	Column A	A
✓ Results	»	6		
Normality and Lognormality Tests of Data 4		7	Mann Whitney test	
🔲 Mann-Whitney test of Data 4  1		8 (	3 P value	0.0163
New Analysis		9	Exact or approximate P value?	Exact
✓ Graphs	>>	10	P value summary	*
📐 Normal QQ plot: Normality and Lognormality Tests of Data 4		11	Significantly different (P < 0.05)	Yes
New Graph		12	One- or two-tailed P value?	Two-tailed
✓ Layouts	>>	13	Sum of ranks in column A,B	2146 , 857
New Layout		14	Mann-Whitney U	316
		15		
		16	Difference between medians	
		17	Median of column A	46.50, n=60
	_	18	Median of column B	70.00, n=17
Family	»	19	Difference: Actual	23.50
📃 Data 4		20	Difference: Hodges-Lehmann	20.00
Mann-Whitney test		21		

#### ① 点击M-W检测显示出右侧检验结果

- ② 点击此按钮可再次调出检验设置窗口,修改检验设置
- ③ *p* 值=0.0163,两组数据具有显著性差异

# 科研技能单元课

### 独立t检验

- 示例: 文献示例: Han J, Wang JZ, Yang X, Yu H, Zhou R, Lu HC, Yuan WB, Lu JC, Zhou ZJ, Lu Q, Wei JF, Yang H. METTL3 promote tumor proliferation of bladder cancer by accelerating pri-miR221/222 maturation in m6A-dependent manner. Mol Cancer. 2019 Jun 22;18(1):110. doi: 10.1186/s12943-019-1036-9. PMID: 31228940; PMCID: PMC6588935. Fig.1a
- 示例解读:
  - ▶ 该图是一个mettl3在正常组织与肿瘤组织中的mRNA含量的对比
  - ▶ 图中X轴仅以不同处理条件作为分类依据(Tumor,Normal),

每个分类下,有几十次实验结果



• 模拟数据	Table format:		Group A	Group B
		Column	Tumor	Normal
		×		
	1	Title	0.01000	0.00960
	2	Title	0.00100	0.00840
	3	Title	0.00100	0.00390
	4	Title	0.00100	0.00340
	5	Title	0.00200	0.00640
	6	Title	0.00250	0.00540
	7	Title	0.00250	0.01460
	8	Title	0.00250	0.00720
	9	Title	0.02500	0.01220
	10	Title	0.00550	0.01220
	11	Title	0.02550	0.00920
	12	Title	0.00500	0.00220
	13	Title	0.02000	0.01410
	14	Title	0.02000	0.02600
	1			

## 科研技能 单元课

www.helixlife.cn,

### 独立t检验

#### • 分析目标

- ✔ 从数据结构上看,是个一维分组,分为两组,分别为肿瘤组与正常人组
- ✔ 从实验设计上看,是独立设计;



# 科研技能单元课

独立t检验

1 新建一个Column的数据表(重命名为"t test"),将数据按下图的格式输入

①选择新建表格;

②输入每个值, excel可以拷贝;

New Data Table and Graph	
	Column tables have one grouping variable, with each group defined by a column
CREATE	A B
XY	Control Treated
Column	
Grouped	2 Control Treated @ Learn more
Contingency	Data table:
Survival	
Parts of whole	Start with sample data to follow a tutorial
Multiple variables	
Nested	Options:
	2 DEnter replicate values, stacked into columns
ODEN	Enter paired or repeated measures data - each subject on a separate row
OPEN	Enter and plot error values already calculated elsewhere
Clone a Graph	Enter: Mean, SD, N $\checkmark$

按前述步骤检验正态分布
 ③点击分析后选择正态检测;
 ④勾选所有数据

Analyze Data	×
Use: Built-in analysis	$\checkmark$
Which analysis?	Analyze which data sets?
Search	Table: Data 1 $$
<ul> <li>Recently used</li> <li>Transform, Normalize Transform Transform concentrations (X) Normalize Prune rows Remove baseline and column math Transpose X and Y Fraction of total</li> <li>XY analyses</li> <li>Column analyses t tests (and nonparametric tests) One-way ANOVA (and nonparametric or One sample t and Wilcoxon test Descriptive statistics</li> <li>Normality and Lognormality Tests Frequency distribution ROC Curve</li> </ul>	A:Tumor B:Normal

## 科研技能 单元课

①选择T检验

②勾选所有数据

独立t检验

#### ✔ 通过正态检验(P value均大于0.1)

<b>\$</b>	Normality and Lognormality Tests	А	В
	Normality and Lognormality rests	Tumor	Normal
1	Test for normal distribution		
2	Anderson-Darling test		
3	A2*	0.5333	0.5251
4	P value	0.1650	0.1731
5	Passed normality test (alpha=0.05)'	Yes	Yes
6	P value summary	ns	ns
7			
8	D'Agostino & Pearson test		
9	K2	1.095	2.152
10	P value	0.5783	0.3409
11	Passed normality test (alpha=0.05)'	Yes	Yes
12	P value summary	ns	ns
13			
14	Shapiro-Wilk test		
15	W	0.9651	0.9571
16	P value	0.1164	0.0513
17	Passed normality test (alpha=0.05)'	Yes	Yes
18	P value summary	ns	ns
19			
20	Kolmogorov-Smirnov test		
21	KS distance	0.09936	0.09648
22	P value	>0.1000	>0.1000
23	Passed normality test (alpha=0.05)'	Yes	Yes
24	P value summary	ns	ns

#### 3 回到column数据页面,点击顶部Analyze按钮,弹出的窗口里设置检验方法



## 科研技能 单元课

独立t检验

#### 4 弹出的窗口里设置检验参数,然后点击"OK"

Parameters: t tests (and No	nparametric Tests)	×	Parameters: t tests (and N
Experimental Design Residuals	Options		Experimental Design Residua
Experimental design  Unpaired  Paired  A Control  Y 1 2	B Treated Y		Calculations P value: One-tailed Report differences as: B Confidence level: 955 Definition of statistical si Graphing options Graph differences (pairs Graph ranks (nonparam
3 4 5			Graph correlation (paire Graph CI of difference b Additional results
Assume Gaussian distribution	n?		Descriptive statistics for     t test: Also compare mo     Mann-Whitney: Also con     Assumes both distribut
3 Onpaired t test. Assume b	oth populations have the same SD		Wilcoxon: When both va
O Unpaired t test with Welch	's correction. Do not assume equal SDs		Output Show this many significant
			P value style: GP: 0.1234
			Make options on this tab b
	Learn Cancel	ок	

Parameters: t tests (and Nonparametric Tests)	×				
Experimental Design Residuals Options					
Calculations					
P value: One-tailed  Two-tailed (recommended)					
Report differences as: $B - \Delta$	~				
Confidence level: 95% ~					
Definition of statistical significance: P < 0.05					
Graphing options					
Graph differences (paired)					
Graph ranks (nonparametric)					
Graph correlation (paired)					
✓ Graph CI of difference between medians					
Additional results					
Descriptive statistics for each data set					
t test: Also compare models using AICc					
Mann-Whitney: Also compute the CI of difference between medians					
Assumes both distributions have the same shape.					
Wilcoxon: When both values on a row are identical, use method of Pratt					
If this option is unchecked, those rows are ignored.					
Output					
Show this many significant digits (for everything except P values):					
P value style: GP: 0.1234 (ns), 0.0332 (*), 0.0021 (**), V = 6					
Make options on this tab be the default for future tests.					
Learn Cancel O	К				

①选择非配对设计

②服从正态分布,勾选参数检验

③服从方差齐性,选择非配对t检验不需矫正

不知道数据是否方差齐性时,可先随意选择一个;分析结

果里会显示F检验的结果,据此判断数据是否服从方差齐

性,再来修改这个选项

在大多情况下,选项中不需要再行修改

④点击OK

# 科研技能单元课

(4)

OK

独立t检验

#### う 查看分析结果

Search ~		Tabul ar results 🗸 🗸	
✓ Data Tables »	8	Unpaired t test	
🖽 Data 1		Tabular results	
🕀 New Data Table	-		
✓ Info »	1	Table Analyzed	Data 1
<ol> <li>Project info 1</li> </ol>	2		
New Info	3	Column B	Normal
✓ Results »	4	VS.	VS.
🔳 Unpaired t test of Data 1	5	Column A	Tumor
Normality and Lognormality Te	6		
New Analysis	7	Unpaired t test	
✓ Graphs »	8	P value	0.0049
🗠 Data 1	9	P value summary	**
Estimation Plot: Unpaired t tes	10	Significantly different (P < 0.05)?	Yes
New Graph	11	One- or two-tailed P value?	Two-tailed
v Lavouts »	12	t, df	t=2.872, df=106
Mew Lavout	13		
Vew Layout	14	How big is the difference?	
	15	Mean of column A	0.01584
	16	Mean of column B	0.01132
Family »	17	Difference between means (B - A) ± SEN	-0.004515 ± 0.001572
🖽 Data 1	18	95% confidence interval	-0.007631 to -0.001398
Unpaired t test	19	R squared (eta squared)	0.07221
Estimation Plot: Unpaired t test (	20		
•	21	F test to compare variances	
	22	F, DFn, Dfd	2.393, 53, 53
	23	P value	0.0018
	24	P value summary	**
	25	Significantly different (P < 0.05)?	Yes



Learn

Cancel

 ① p =0.0049,两组有显著性 差异;
 ② F检验,一般要求p>0.1 (也有人认为p >0.05),认
 为方差齐性;这里p <0.1, 可认为两组数据方差不齐, 返回选择;
 ③选择Welch's correlation
 ④点击OK

## 科研技能 单元课

独立t检验

#### 添加显著性差异

1	Welch's t test • 矫正后	的结果
1	Table Analyzed	Data 1
2		
3	Column B	Normal
4	VS.	VS.
5	Column A	Tumor
6		
7	Unpaired t test with Welch's correction	
8	P value	0.0051 (1)
9	P value summary	**
10	Significantly different (P < 0.05)?	Yes
11	One- or two-tailed P value?	Two-tailed
12	Welch-corrected t, df	t=2.872, df=90.71 (2)
13		
14	How big is the difference?	
15	Mean of column A	0.01584
16	Mean of column B	0.01132
17	Difference between means (B - A) ± SEM	-0.004515 ± 0.001572
18	95% confidence interval	-0.007637 to -0.001392
19	R squared (eta squared)	0.08337
20		
21	F test to compare variances	
22	F, DFn, Dfd	2.393, 53, 53
23	P value	0.0018
24	P value summary	**
25	Significantly different (P < 0.05)?	Yes
26		



①*p*=0.0051,两组有显著性差异;
②矫正后;
③点击左侧工具栏Graph出现图像,点击此处, 自动添加显著性差异符号(点击还可以对其进行 格式修改。自动添加结果显示如图中所示)

# 科研技能 单元课

### 配对t检验

- 对于配对设计的实验,就需要进行配对T检验
- 示例文献: Jafarimanesh H, Vakilian K, Mobasseri S. Thermo-therapy and cryotherapy to decrease the symptoms of restless leg syndrome during the pregnancy: A randomized clinical trial. Complement Ther Med. 2020 May;50:102409. doi: 10.1016/j.ctim.2020.102409. Epub 2020 Apr 19. PMID: 32444058. Table.2

模拟

• 示例解读:

▶ 在37名孕妇,以同一人水疗前后的不宁腿指数作为两组,衡量治疗的效果

Table 2Mean of RLS in the intervention groups.				
Cold Mean ± SD	Warm Mean ± SD			
$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	$\begin{array}{rrrr} 19.32 \ \pm \ 5.87 \\ 13.50 \ \pm \ 4.74 \\ 0.001 \end{array}$			
	Cold Mean ± SD 18.65 ± 6.16 11.02 ± 4.93 0.001			

*++-			
釵掂	-#	Group A	Group B
		Data Set-A	Data Set-B
	1	31.7	14.5
	2	16.9	15.2
	3	18.8	15.6
	4	20.0	12.9
	5	27.1	15.7
	6	17.1	9.4
	7	21.9	7.1
	8	31.3	2.4
	9	20.0	3.8
	10	30.8	16.3
	11	16.9	16.6
	12	19.8	15.8
	13	12.3	7.5
	14	21.4	11.2
	15	17.9	12.7
	16	4.6	11.7
	17	18.7	14.0

## 科研技能 单元课

配对t检验

#### • 分析目标

✓数据大体跟上一例差不多,不同的地方在于,本例中的两组数据,是配对关系,因此需要使用配对t检验。



# 科研技能单元课

配对t检验

1 先在excel里计算两组数据的差值

J	K	L	1
Before	After	After-before	• =14 5-31 7
31.7	14.5	-17.2	• =15.2-16.9
16.9	15.2	-1.7	
18.8	15.6	-3.2	
20.0	12.9	-7.1	
27.1	15.7	-11.4	
17.1	9.4	-7.6	
21.9	7.1	-14.8	
31.3	2.4	-28.9	
20.0	3.8	-16.2	
30.8	16.3	-14.5	
16.9	16.6	-0.3	
19.8	15.8	-4.1	
12.3	7.5	-4.8	
21.4	11.2	-10.2	
17.9	12.7	-5.1	
4.6	11.7	7.0	
18.7	14.0	-4.7	
22.2	10.8	-11.4	
12.6	4.6	-8.0	
15.2	8.1	-7.1	



Welcome to GraphPad Prism			×
GraphPad Prism Version 9.3.1 (471)	Column tables have one grouping variable, with each group defined by a column		
CREATE	Data table:		
XY Column 2 Grouped	Enter or import data into a new table     Start with sample data to follow a tutorial		
Contingency	Options:		
Survival	Enter replicate values, stacked into columns		
Parts of whole	Enter paired or repeated measures data - each subject on a separate row		
Nested	Enter: Mean, SD, N		
LEARN			
Getting Started			
Videos 🔗			
Prism Guides			
Graph Portfolio			
OPEN			
Open a file			
LabArchives			
Clone a Graph			
			5
Prism Tips		Cancel	Create

①在excel中利用公式计算
 After-before
 ②新建Column
 ③创建新table
 ④选择输入每个值
 ⑤点击OK

## 科研技能 单元课

配对t检验



1	Group A	Group B	Group C
	Before	After	After-before
	Y	Y	Y
1	31.7	14.5	-17.2
2	16.9	15.2	-1.7
3	18.8	15.6	-3.2
4	20.0	12.9	-7.1
5	27.1	15.7	-11.4
6	17.1	9.4	-7.6
7	21.9	7.1	-14.8
8	31.3	2.4	-28.9
9	20.0	3.8	-16.2
10	30.8	16.3	-14.5
11	16.9	16.6	-0.3
12	19.8	15.8	-4.1
13	12.3	7.5	-4.8
14	21.4	11.2	-10.2
15	17.9	12.7	-5.1
16	4.6	11.7	7.0
17	18.7	14.0	-4.7

#### 按前述步骤,对两组的差值进行正态分布检验



①excel拷贝数据 ②点击分析后,选择正态检测 ③只对差值进行正态检测 ④四种检测方法显示符合正态分布 四种检测方法结果显示P值均大于0.1

(4)	Test for normal distribution		15		
2	Anderson-Darling test		14	Shapiro-Wilk test	
~	Anderson-Darning test		15	W	0.9775
3	A2*	0.2536	16	P value	0 6445
4	P value	0.7142	10		0.0440
5	Passed normality test (alpha=0.05)?	Yes	17	Passed normality test (alpha=0.05)?	Yes
0		100	18	P value summary	ns
6	P value summary	ns	19		
7			00		
8	D'Agostino & Pearson test		20	Kolmogorov-Smirnov test	
0	K2	1 366	21	KS distance	0.07687
9		1.300	22	P value	>0.1000
10	P value	0.5052	22	Passed pormality test (alpha=0.05)2	Vec
11	Passed normality test (alpha=0.05)?	Yes	23		163
12	P value summary	ne	24	P value summary	ns
12	r value summary	113	25		

## 科研技能 单元课

#### 配对t检验



Analyze Data			×
Use: Bui	ilt-in analysis	$\sim$	
Which analysis?		Analyze which data sets?	
Search		Table: pair t test	$\sim$
Recently used	^	A:Before	
<ul> <li>□ Transform, Normalize         Transform         Transform concentrations (X)         Normalize         Prune rows         Remove baseline and column m         Transpose X and Y         Fraction of total         # XY analyses         Column analyses         Column analyses         (1) t tests (and nonparametric test             One-way ANOVA (and nonpara             One sample t and Wilcoxon test             Descriptive statistics             Normality and Lognormality Test             Frequency distribution             ROC Curve             Bland-Altman method comparise             Identify outliers             Analyze a stack of P values             # Grouped analyses         </li> </ul>	nath metric or t sts	B:After C:After-before	
Contingency table analyses	~		_
<	>	Select All Deselect All	
		Help Cancel 0	ĸ



rvhei	imental	Design Residuals I design	Options		
0	Unpaired	1			
$\bigcirc$	Paired	]4			
		Α	В	3	
		Control	Treated	3	
		Y	Y	~	
	1			3	
	2			3	
	3				
	4			3	
	5				
	20			2	
_		$\cdots$		2	
Assu	me Gau	ssian distributio	n?	~	
Assu	me Gaus Yes. Use	ssian distributio	n? 5	N.	
Assur	me Gaus Yes. Use No. Use	ssian distributio parametric test. nonparametric tes	n? 5	~	
Assur Choos	me Gaus Yes. Use No. Use se test	ssian distributio parametric test. nonparametric tes	n? 5_ st.	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	
Assur	me Gaus Yes. Use No. Use se test Paired t	ssian distributio parametric test. nonparametric tes test (differences b	n? 5.	lues are consistent) (	6
Assur Choose	me Gaus Yes. Use No. Use Se test Paired t Ratio pai	ssian distributio parametric test. nonparametric test test (differences b ired t test (ratios o	n? 5. 	lues are consistent) ( re consistent)	6
Assur Choose	me Gaus Yes. Use No. Use se test Paired t Ratio pai	ssian distributio parametric test. nonparametric test test (differences b ired t test (ratios o	n? 5 .st. Detween paired va of paired values a	lues are consistent) ( re consistent)	6
Assur Choose	me Gaus Yes. Use No. Use se test Paired t Ratio pa	ssian distributio parametric test. nonparametric test test (differences b ired t test (ratios o	n? 5.	lues are consistent) ( re consistent)	6
Assur Choos	me Gaus Yes. Use No. Use se test Paired t Ratio pai	ssian distributio parametric test. nonparametric tes test (differences b ired t test (ratios o	n? 5 st. Detween paired va of paired values a	lues are consistent) ( re consistent)	6
Assur Choo:	me Gaus Yes. Use No. Use se test Paired t Ratio pai	ssian distributio parametric test. nonparametric test test (differences b ired t test (ratios o	n? 5 st.	lues are consistent) ( re consistent)	6
Assur Choos	me Gaus Yes, Use No. Use se test Paired t Ratio pai	ssian distributio parametric test. nonparametric test test (differences b ired t test (ratios o	n? 5. between paired va of paired values a	lues are consistent) ( re consistent)	6)

①选择t检验
 ②勾选差值组
 ③点击OK
 ④勾选配对设计
 ⑤服从正态分布
 ⑥选择配对t检验
 ⑦点击OK

# 科研技能单元课

配对t检验

#### 查看分析结果

Search ~	Tabular results 🗸						
✓ Data Tables »	-	Paired t test					
🧮 pair t test		Tabular results					
🕀 New Data Table							
✓ Info »	1	Table Analyzed	pair t test				
<ol> <li>Project info 1</li> </ol>	2						
New Info	3	Column B	After				
✓ Results »	4	VS.	VS.				
Normality and Lognormality Te	5	Column A	Before				
📄 Paired t test of pair t test	6		<u> </u>				
New Analysis	7	Paired t test					
✓ Graphs »	8	P value	<0.0001				
🕅 pair t test	9	P value summary	****				
Normal OO plot: Normality and	10	Significantly different (P < 0.05)?	Yes				
Estimation Plot: Paired t test o	11	One- or two-tailed P value?	Two-tailed				
New Graph	12	t, df	t=6.367, df=36				
v Lavouts	13	Number of pairs	37				
New avout	14						
• New Layout	15	How big is the difference?					
	16	Mean of differences (B - A)	-7.878				
Family »	17	SD of differences	7.527				
📃 pair t test	18	SEM of differences	1.237				
Paired t test	19	95% confidence interval	-10.39 to -5.369				
Estimation Plot: Paired t test of r	20	R squared (partial eta squared)	0.5296				

#### 8 显著性差异标注

Format Pairwise Comparisons X	显著性差异
Appearance Comparisons on Graph	单击自动标
Display options 1	行修改
P value (numbers)	
OAsterisks representing P value classification	
P value threshold (2)	
Display asterisks/values and lines for which P values	①展示方式
All P values (including "ns")	
$\bigcirc$ P values less than or equal to 0.05 $\checkmark$	
Line/bracket and text options	③点的形式
Plot: 3	④线厚度
Thickness: 1 pt V	
Color: 5	⑤颜色
Style: (6)	
Asterisks or value: Font Arial, 12pt	6)风格
Use the options on this tab as default	
Help Cancel Apply OK	

### 显著性差异的做法就不再赘述 单击自动标注可对标注形式进 行修改

# 科研技能单元课

### 方差分析

- 示例: 文献Liang Y, Song X, Li Y, Chen B, Zhao W, Wang L, Zhang H, Liu Y, Han D, Zhang N, Ma T, Wang Y, Ye F, Luo D, Li X, Yang Q. LncRNA BCRT1 promotes breast cancer progression by targeting miR-1303/PTBP3 axis. Mol Cancer. 2020 May 8;19(1):85. doi: 10.1186/s12943-020-01206-5. PMID: 32384893; PMCID: PMC7206728. Fig.3C
- 示例解读:
  - ▶ 该图是一个小管形成实验,一般是在内皮细胞做实验,论证血管新生能力的大小
  - ▶ 图中X轴仅以不同处理条件作为分类依据(si-NC, si-BCRT1, pcDNA3.1, lncRNA BCRT1),每 种分类下,有几次实验结果



• 模拟数据

si-NC	si-BCRT1	pcDNA3.1	IncRNA BCRT1
1.01	0.44	1.11	3.58
1.11	0.45	1.02	3.16
0.95	0.35	0.95	3.94

## 科研技能 单元课

### 方差分析

#### • 分析目标

- ✔ 这是一个连续性的数据,以均值的形式表现,每组数据样本量少,经正态检测后是正态分布,符合方差分析的选择;
- ✓ GraphPad Prism中,各组数据服从正态分布是使用One-way ANOVA(单因素方差分析)的前提,若数据不服从正态分布,有几个选择:①通过转换使得数据服从 正态分布,如log、倒数等;②使用非参数检验;③直接进行分析:由于单因素方差分析对于偏离正态分布比较稳健,尤其是在各组样本量相等或近似相等的情况 下,而且非正态分布实质上并不影响犯Ⅰ型错误的概率



## 科研技能 单元课

方差分析



Analyze Data	×
Use: Built-in analysis	~
Which analysis?	Analyze which data sets?
Search	Table: Data 1
<ul> <li>Recently used</li> <li>Transform, Normalize Transform</li> <li>Transform concentrations (X) Normalize</li> <li>Prune rows</li> <li>Remove baseline and column math</li> <li>Transpose X and Y</li> <li>Fraction of total</li> <li>XY analyses</li> <li>Column analyses</li> <li>t tests (and nonparametric tests)</li> <li>One-way ANOVA (and nonparametric or One sample t and Wilcoxon test</li> <li>Descriptive statistics</li> <li>Normality and Lognormality Tests</li> <li>Frequency distribution</li> <li>ROC Curve</li> <li>Bland-Altman method comparison</li> <li>Identify outliers</li> <li>Analyze a stack of P values</li> </ul>	A:si-NC B:si-BCRT1 C:pcDNA3.1 D:lncRNA BCRT1 3
	Select All Deselect All
	Help Cancel OK

2 在弹出的窗口中,勾选正态分布检验的方法,点击"OK"



## 科研技能 单元课

www.helixlife.cn/

### 方差分析

### 3 查看结果

▶ 正态分布检验时,一般认为,

当p value > 0.1 时,数据服从正态分布

- ▶ 由于样本量过小,只有S-W检测能够有效
- ▶ 四组数据均符合正态分布

8	Normality and Lognormality Tests	Α	В	С	D
	Tabular results	si-NC	si-BCRT1	pcDNA3.1	IncRNA BCRT1
1	Test for normal distribution				
2	Anderson-Darling test				
3	A2*	N too small	N too small	N too small	N too small
4	P value				
5	Passed normality test (alpha=0.05)?				
6	P value summary				
7					
8	D'Agostino & Pearson test				
9	К2	N too small	N too small	N too small	N too small
10	P value				
11	Passed normality test (alpha=0.05)?				
12	P value summary				
13					
14	Shapiro-Wilk test				
15	W	0.9796	0.8242	0.9948	0.9980
16	P value	0.7262	0.1736	0.8624	0.9152
17	Passed normality test (alpha=0.05)?	Yes	Yes	Yes	Yes
18	P value summary	ns	ns	ns	ns
19		• <i>P</i> 值均	大于0.1、服/	人正态分布	
20	Kolmogorov-Smirnov test				
21	KS distance	N too small	N too small	N too small	N too small
22	P value				
23	Passed normality test (alpha=0.05)?				
24	P value summary				
25					
26	Number of values	3	3	3	3

## 科研技能 单元课

www.helixlife.cn

### 方差分析

### 4 分析数据

#### ①点击Analyze

②在Column analyses下找到One-way ANOVA

③默认已勾选了所有组别

#### ④点击OK

Clip	board Ar	oard Analysis		Change	Import	Draw		
& 1	Ê	/ze 🎽 🎢	EI E	Ì ≩I 2↓ - &	• •	txt xml	* •	
	Group A	Group E	}	Group C		Group D		
	si-NC	si-BCRT	1 pcDNA3.1		In	cRNA B	CRT1	
1	1.01	0.	44	1.11			3.58	
2	1.11	0.	45	1.02			3.16	
3	0.95	0.	35	0.95			3.94	





## 科研技能 单元课

### 方差分析

#### 6 Experimental Design

Param	eters:	One-Way ANC	VA (and Nonp	arametric or I	Vixed)	>				
Exper	Experimental Design Repeated Measures Multiple Comparisons Options Residuals									
Expe	rimenta	l design								
	No matching or pairing									
Cach row represents matched, or repeated measures, data										
		Group A	Group B	Group C	Group D	> 5				
		Data Set-A	Data Set-B	Data Set-C	Title					
		Y	Y	Y	Y	3				
	1					5				
	2	$\bigcirc$	$\bigcirc$	$\mathbf{O}$		3				
	~~		him	~~~~	h	20				
Assu	me Gau	ssian distributio	n of residuals?							
	Yes. Use	ANOVA.	)							
0	No. Use	nonparametric te	st.							
Assu	me equ	al SDs?								
0	Yes. Use	ordinary ANOVA	test.		(4)					
0	No. Use	Brown-Forsythe a	nd Welch ANOVA	tests.						
_										
Bas	sed on yo	our choices (on all	tabs), Prism will	perform:						
	- Ordinar	γ one-way ANOV	Α.							
			Lear	n Cai	ncel	ОК				

#### Multiple Comparisons

Parameters: One-Way ANOVA (and Nonparametric or Mixed)	<
Experimental Design Repeated Measures Multiple Comparisons Options Residuals	
Followup tests	
○ None.	
• Compare the mean of each column with the mean of every other column.	
$\bigcirc$ Compare the mean of each column with the mean of a control column.	
Control column: Column A : si-NC ~	
$\bigcirc$ Compare the means of preselected pairs of columns.	
Selected pairs: Select	
$\bigcirc$ Test for linear trend between column mean and left-to-right column order.	
Which test?	
Use choices on the Options tab to choose the test, and to set the defaults for future ANOVAs.	
Learn Cancel OK	ה

①点击 Experimental Design
 ②选择非重复或非配对测量
 ③服从正态分布,选择ANOVA
 ④方差齐,选择经典ANOVA
 ⑤选择Multiple Comparisons
 ⑥选择两两比较
 ⑦选择OK

### 科研技能 单元课 www.helixlife.cn/

方差分析

#### 查看结果 8

		(2)						
Search ~	Je	ANOVA results × 🔲 Multiple comparisons	$\times   \sim  $					① 在左边导航栏甲找到"Ordinary one-
✓ Data Tables »	8							
🖽 Data 1		ANOVA results lits						M(2)/
🕀 New Data Table								way ···
✓ Info »	4	ANOVA summary						
<ol> <li>Project info 1</li> </ol>	5	F	139.6					」(2) 点击具目录下的 "ANOVA" 查看结果
New Info	6	P value	<0.0001		3			
✓ Results »	7	P value summary	****					」 ③ 统计结果: p 值的意义和p 值。这里
Normality and Lognormality Te	8	Significant diff. among means (P < 0.05)'	Yes					
1 Ordinary one-way ANOVA of	9	R squared	0.9813					值<0.05表示:多组数据中,至少有
New Analysis	10							
▼ Graphs »	11	Brown-Forsythe test						
M Data 1	12	F (DFn, DFd)	2.340 (3, 8)					
Normal OO plot: Normality and	13	P value	0.1496					
New Graph	14	P value summary	ns					
v Lavouts »	15	Are SDs significantly different (P < 0.05)?	No					
Alex Lavout	16							差异,则需要查看多重比较的结果
Wew Layout	17	Bartlett's test						
	18	Bartlett's statistic (corrected)						
	19	P value						
Family »	20	P value summary						
Data 1	21	Are SDs significantly different (P < 0.05)						
Ordinary one-way ANOVA	22							
	23	ANOVA table	SS	DF	MS	F (DFn, DFd)	P value	
	24	Treatment (between columns)	17.63	3	5.876	F (3, 8) = 139.6	P<0.0001	
	25	Residual (within columns)	0.3368	8	0.04210			
	26	Total	17.96	11				

「的"ANOVA"查看结果 值的意义和*p* 值。这里*p* : 多组数据中,至少有一 E) 或者多对具有显著性差 本是哪个组跟哪个组之间有 查看多重比较的结果

## 科研技能 单元课

www.helixlife.cn/

### 方差分析

#### 8 查看结果

- ① 在左边导航栏里找到"Ordinary one-way ANOVA"
- ② 点击其目录下的 "Multiple comparisons" 查看结果 ( "ANOVA "表格里是总体统计结果)

 $\bigcirc$ 

③ 统计结果: p值的意义和p值。第6-11行分别显示各组两两比较的结果

Search ~	E	ANOVA results 🗙 📄 <b>Hultiple comparis</b> d	ns 🗙 🗸 🗸							
✓ Data Tables »	8	Ordinary one-way ANOVA								
🖽 Data 1		Multiple comparisons								
🕀 New Data Table		1								
✓ Info »	1	Number of families	1							
<ol> <li>Project info 1</li> </ol>	2	Number of comparisons per family	6							
New Info	3	Alpha	0.05							
✓ Results >>	4	3							$\mathcal{A}$	
Normality and Lognormality Te	5	Tukey's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Below threshold?	Summary	Adjusted P Value			
Ordinary one-way ANOVA of	6	si-NC vs. si-BCRT1	0.6100	0.07351 to 1.146	Yes	*	0.0271	A-B		
New Analysis	7	si-NC vs. pcDNA3.1	-0.003333	-0.5398 to 0.5332	No	ns	>0.9999	A-C		
<ul> <li>Graphs</li> </ul>	8	si-NC vs. IncRNA BCRT1	-2.537	-3.073 to -2.000	Yes	****	<0.0001	A-D		
N Data 1	9	si-BCRT1 vs. pcDNA3.1	-0.6133	-1.150 to -0.07684	Yes	*	0.0264	B-C		
Normal OO plati Normality and	10	si-BCRT1 vs. IncRNA BCRT1	-3.147	-3.683 to -2.610	Yes	****	<0.0001	B-D		
	11	pcDNA3.1 vs. IncRNA BCRT1	-2.533	-3.070 to -1.997	Yes	****	<0.0001	C-D		
New Graph	12								/	
✓ Layouts ×	13	Test details	Mean 1	Mean 2	Mean Diff.	SE of diff.	n1	n2	q	DF
I New Layout	14	si-NC vs. si-BCRT1	1.023	0.4133	0.6100	0.1675	3	3	5.149	8
	15	si-NC vs. pcDNA3.1	1.023	1.027	-0.003333	0.1675	3	3	0.02814	8
	16	si-NC vs. IncRNA BCRT1	1.023	3.560	-2.537	0.1675	3	3	21.41	8
Family »	17	si-BCRT1 vs. pcDNA3.1	0.4133	1.027	-0.6133	0.1675	3	3	5.177	8
🖽 Data 1	18	si-BCRT1 vs. IncRNA BCRT1	0.4133	3.560	-3.147	0.1675	3	3	26.56	8
Ordinary one-way ANOVA	19	pcDNA3.1 vs. IncRNA BCRT1	1.027	3.560	-2.533	0.1675	3	3	21.39	8

## 科研技能 单元课

### 方差分析



<u>*</u> • √a <u>w</u> <u>i</u>							
Automatically add pairwise comparisons							
Format pairwise comparisons							
2) Choose pairwise comparisons to plot							
Remove pairwise comparisons							
Manually add lines with text							
Format Pairwise Comparisons	$\times$						
Appearance Comparisons on Graph							
Appearance companions on Graph							
Ordinary one-way ANOVA of Data 1	~						
Select comparisons to be displayed on the graph	_						
si-NC vs. si-BCRT1							
si-NC vs. pcDNA3.1							
si-BCRT1 vs. pcDNA3.1							
si-BCRT1 vs. IncRNA BCRT1							
pcDNA3.1 vs. IncRNA BCRT1							
Select All Deselect All							
Help Cancel Apply OK							

- ① 点击可自动出现所有的comparisons,如左侧图
- ② 点击三角形下拉选项,选择第三选项
- ③ 将不需要的选项勾去

#### ④ 结果如下图



## 科研技能 单元课

### 方差分析

• 示例: 文献Liang Y, Song X, Li Y, Chen B, Zhao W, Wang L, Zhang H, Liu Y, Han D, Zhang N, Ma T, Wang Y, Ye F, Luo D, Li X, Yang Q. LncRNA BCRT1 promotes breast cancer progression by targeting miR-1303/PTBP3 axis. Mol Cancer. 2020 May 8;19(1):85. doi: 10.1186/s12943-020-01206-5. PMID: 32384893; PMCID: PMC7206728. Fig.2b

#### • 示例解读:

▶ 2个组别,分别是 si-NC, si-BCRT1;每个组 别检测了0、1、2、3、4、5、6天的数据, 每个数据有3个重复

• 模拟数据



Table	format:	Х		Group A		Group B			
>	(Y	day		si-NC		si-BCRT1			
-	×	Х	A:Y1	A:Y2	A:Y3	B:Y1	B:Y2	B:Y3	
1	Title	1.0	1.0000	1.0000	1.0000	1.0	1.0	1.0	
2	Title	2.0	1.2000	1.6000	1.8000	2.0	2.3	2.4	
3	Title	3.0	2.5000	2.6000	2.7000	4.3	4.4	4.6	
4	Title	4.0	3.9000	3.7000	3.5000	5.1	5.2	5.4	
5	Title	5.0	4.0000	4.1000	4.5000	5.4	5.7	5.9	
6	Title	6.0	5.5000	5.6000	5.9000	6.2	6.3	6.5	

## 科研技能 单元课

方差分析

#### • 分析目标

- ✔ 数据结构上,这是个二维分组,处理分组与时间点
- ✔ 从实验结果上看,需要比较实验组和对照组相同时间点上的数据是否有差异,细胞间不需要比较,但需要比较的时间点有多个,双因素,用Two-way ANOVA



## 科研技能 单元课

### 方差分析



#### ①excel拷贝数据

#### ②点击分析后

③勾选Grouped analyses下的Two-way ANOVA

④选择所有组别

⑤选择OK

Fa	mily Windo	w Help							
	Analysis	Cha	nge Imp	ort Draw	۷	Vrite			Text
/		<b>≤</b> ∎ <b>≥</b> ∎ A <sub>z</sub> ,	l- 🖄 - 🦽	* -	√ā	w 🗓	~		~ <u>A</u>
	Analyze	<i><sup>*</sup></i> <sup>12</sup> <sup>±</sup> <sup>±</sup>	<sup>3</sup> <del>1.23</del> ()*-	•	T	T X-	A A	• <b>B</b> <i>I</i> <u>U</u> 2	X² X₂ ₪ ₪
	X		Group A					Group B	
	day	1	si-NC					si-BCRT1	
	Х	A:Y1	A:Y2	A:Y3	3	B:)	(1	B:Y2	B:Y3
	1.0	1.0000	1.0000	1.0	0000		1.0	1.0	1.0
	2.0	1.2000	1.6000	1.8	3000		2.0	2.3	2.4
	3.0	2.5000	2.6000	2.7	7000		4.3	4.4	4.6
	4.0	3.9000	3.7000	3.5	5000		5.1	5.2	5.4
	5.0	4.0000	4.1000	4.5	5000		5.4	5.7	5.9
	6.0	5.5000	5.6000	5.9	000		6.2	6.3	6.5

Analyze Data	×
Use: Built-in analysis	$\sim$
Which analysis?	Analyze which data sets?
Search	Table: Data 1 $\lor$
<ul> <li>Column analyses         <ul> <li>t tests (and nonparametric tests)</li> <li>One-way ANOVA (and nonparametric or One sample t and Wilcoxon test Descriptive statistics</li> <li>Normality and Lognormality Tests</li> <li>Frequency distribution</li> <li>ROC Curve</li> <li>Bland-Altman method comparison Identify outliers</li> <li>Analyze a stack of P values</li> </ul> </li> <li>Grouped analyses</li> <li>Two-way ANOVA (or mixed model)</li> <li>Row statistics</li> <li>Multiple t tests (and nonparametric tests</li> <li>Contingency table analyses</li> <li>Survival analyses</li> <li>Parts of whole analyses</li> <li>Multiple variable analyses</li> <li>Generate curve</li> <li>Simulate data</li> </ul>	A:si-NC B:si-BCRT1
	(5) Help Cancel OK

## 科研技能 单元课

方差分析

#### 2 选择实验参数

Table format: Group A Group B Group C							
Grouped		Title		Ti	tle	Ti	tle
		A:Y1	A:Y2	B:Y1	B:Y2	C:Y1	C:Y2
1	Time1						
2	Time2						
3	Time3						
4	Time4	han	Low	han	m	$\sim$	m
Eac ude	h column rep h row repres <b>interaction</b> Fit a main ef	resents a dif ents a differ <b>term?</b> fects only m	fferent time poin ent time poin odel (column	ooint, so mato t, so matche effect and ro	ched values a d values are ow effect onl	re spread ac stacked into a y).	rross a row. a subcolumn
] Eac ude ) No. ) Yes ume	h column rep h row repres interaction Fit a main ef Fit a full mo sphericity Use the Geis	resents a dif ents a differ <b>term?</b> fects only m del (column <b>(equal vari</b> ser-Greenho	fferent time p ent time poin odel (column effect, row e <b>ability of dif</b> use correctio	effect and ro effect and ro effect, and co fferences)?	thed values and d values are now effect only lumn/row int nded.	ire spread ac stacked into a γ). eraction effe	cross a row. a subcolumn

Pa	rameters: Two-Way ANOVA (or Mixed Model)	$\times$
N	40del Repeated Measures 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	
	Learn Cancel OK	

①选择Model

②选择行方向上是不同的时间点(重复测量堆叠在亚列

内)

③选择Repeated Measures

④选择分析方法

⑤选择OK

其它选项基于需要进行更改,一般情况下为默认选项

## 科研技能 单元课

Ŧ

方差分析

#### 2 选择实验参数2



#### Which test?



- ① 选择Multiple Comparisons
- ② 选择下拉菜单第四项

#### ③ 选择OK

Compare each cell mean with the other cell mean in that row
No multiple comparisons
Compare row means (main row effect)
Within each column, compare rows (simple effects within columns)
Compare each cell mean with the other cell mean in that row

- I. 无多重比较,即仅比较列间各组间总体上的差异
- II. 在行间作多重比较,选择此项,可在下方选择是进行两两比较还是设定一个参照, 其他行都跟这个参照比
- Ⅲ. 组内行间比较,即每组内进行行间多重比较
- IV. 每行进行组别比较,即在每一行里的组别间进行比较
- V. 比较方式的示意图
- VI. 两两比较
- VII. 设定一个参照,其他行都跟这个参照比
- 本例只有两个组别,因此不需要设置列间的多重比较

# 科研技能单元课

方差分析

### 3 查看结果

Search	1) 🖬	ANOVA results ×	$\times$ $\vee$					
✓ Data Tables	»	2way ANOVA						① 在左边导航栏里找到"2way ANOVA of
🧮 Data 1		ANOVA results						
🖽 Data 2								Data1",点击其目录下的"ANOVA
🖽 Data 3	1	Table Analyzed	Data 1					
🕀 New Data Table	2							rocults" 本丢结甲
✓ Info	» 3	Two-way RM ANOVA	Matching: Stacked					TESUILS 旦伯纪木
<ol> <li>Project info 1</li> </ol>	4	Assume sphericity?	Yes					
New Info	5	Alpha	0.05 (2)		3			② 谷变量对总体万差变异的贡献
✓ Results	» 6	(						
1 = 2way ANOVA of Data 1	7	Source of Variation	% of total variation	P value	P value summary	Significant?		③ 差异性分析结果: ρ值的意义和ρ值。
New Analysis	8	Time x Column Factor	2.939	<0.0001	****	Yes		
<ul> <li>Graphs</li> </ul>	» 9	Time	88.51	<0.0001	****	Yes		
Data 1	10	Column Factor	7.859	0.0006	***	Yes		区主P值~0.03农小、区支里中主少有一
Data 2	11	Subject	0.3227	0.0105	*	Yes		
Data 3	12							对或者多对数据具有显著性差异,至于
New Graph	13	ANOVA table	SS	DF	MS	F (DFn, DFd)	P value	
	14	Time x Column Factor	3.582	5	0.7163	F (5, 20) = 31.91	P<0.0001	且体是哪两两数据之间有差异,则需要
	15	Time	107.9	5	21.58	F (5, 20) = 961.1	P<0.0001	
	16	Column Factor	9.579	1	9.579	F (1, 4) = 97.41	P=0.0006	本毛々手以拉的付田
Family	» 17	Subject	0.3934	4	0.09834	F (4, 20) = 4.381	P=0.0105	查看多里比牧的结果
📃 Data 1	18	Residual	0.4490	20	0.02245			
🗐 2way ANOVA	19							
	20	Difference between column means						
	21	Mean of si-NC	3.117					
	22	Mean of si-BCRT1	4.148					
	23	Difference between means	-1.032					
	24	SE of difference	0.1045					
	25	95% CI of difference	-1.322 to -0.7414					

## 科研技能 单元课

方差分析

### 3 查看结果

- ① 点击左边导航栏里 "2way ANOVA of Data1"
- ② 点击 "Multiple comparisons" 查看多重比较的结果
- ③ 多重比较的p值及其意义。"Row1"-"Row6"的p值分别对应于day1-day6,两组数据的显著性检验结果

Search ~		🚍 ANOVA results 🗙 📄 <b>Tultiple comparisons</b> 🗙 📀					
<ul> <li>Data Tables</li> </ul>	» 🐅	2way ANOVA					
🖽 Data 1		Multiple comparisons					
🖽 Data 2		4					
🖽 Data 3	1	Compare each cell mean with the other cell mean in that rov					
🕀 New Data Table	2	2					
v Info	» 3	Number of families	1				
(1) Project info 1	4	Number of comparisons per family	6				
New Info	5	i Alpha	0.05		3		
Results (1)	» 6				3		
2way ANOVA of Data 1	7	Šídák's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Below threshold?	Summary	Adjusted P Value
New Analysis	8						
Graphs	» 9	si-NC - si-BCRT1					
Data 1	1	0 Row 1	0.000	-0.4384 to 0.4384	No	ns	>0.9999
Data 2	1	1 Row 2	-0.7000	-1.138 to -0.2616	Yes	***	0.0007
Data 2	1	2 Row 3	-1.833	-2.272 to -1.395	Yes	****	<0.0001
Alow Graph	1	3 Row 4	-1.533	-1.972 to -1.095	Yes	****	<0.0001
Lavouts	1	4 Row 5	-1.460	-1.898 to -1.022	Yes	***	<0.0001
	1	5 Row 6	-0.6633	-1.102 to -0.2249	Yes	**	0.0013

# 科研技能单元课

### 卡方检验

- 示例: Steering Committee of the Physicians' Health Study Research Group. Preliminary report: Findings from the aspirin component of the ongoing Physicians' Health Study. N Engl J Med. 1988 Jan 28;318(4):262-4. doi: 10.1056/NEJM198801283180431. PMID: 3275899. Table1
- 示例解读:
  - ▶ 这是一项前瞻性研究。随机分配给受试者的两种治疗方法:安慰剂或阿司匹林。两组终点:心梗与未发生心梗。这些值是每个类别中的数目

Table	1.	Cardiovasc	ular E	nd F	Points	in	the	Aspirin	Arm	of	the
Phy	ysi	cians' Health	n Stud	y, A	ccordi	ng	to T	reatmen	t Gro	up.	

END POINT	Aspirin	PLACEBO	Relative Risk	95% Confidence Interval	P VALUE
Myocardial infarction					
Fatal	5	18	0.25	0.11-0.56	0.006
Nonfatal	99	171	0.56	0.44-0.71	< 0.00001
Total	104	189	0.53	0.42-0.67	< 0.00001
Stroke					
Fatal	6	2	3.00	0.75-11.98	0.16
Nonfatal	74	68	1.09	0.78-1.52	0.61
Total	80	70	1.15	0.84-1.58	0.41

- 模拟数据
  - ▶ 结果分为两组,分别是心梗与非心梗,治疗为使用安慰剂与使用阿司匹林

Та	able format:	Outcome A	Outcome B
Contingency		Myocardial Infarction	No MI
	×		
1	Placebo	189	10845
2	Aspirin	104	10933

## 科研技能 单元课

www.helixlife.cn/

### 卡方检验

#### • 分析目标

- ✔ 评估心肌梗死发病率之间的差异是否大于预期的偶然性
- ✔ 量化相对风险及其95%置信区间



## 科研技能 单元课

www.helixlife.cn/

### 卡方检验

1 选择分析方法

①excel拷贝数据

②点击分析后

③从列联表分析列表中选择"Chi-square (and Fisher's exact) test" ④选择所有组别

⑤选择OK

Clipboard	A	nalysis	Cha	nge	Impo
8 🖣	$\chi^2$ %	2	E∎ <b>≥</b> ∎ Az	↓ • <u></u> •	The second
₿ ₿-	Anal	yze 🛅 🎢	<b>1</b> 2: #.# ↓	<sup>3</sup> <del>1.23</del> 💣 -	xml
Table fo	rmat: 1	Outco	me A	Outcom	e B
Conting	ency	Myocardia	Infarction	No M	I
	×				
1 Place	bo		189	10	845
2 Aspir	in		104	10	933

Analyze Data	$\times$
Use: Built-in analysis $\sim$	
Which analysis? Analyze which data sets?	
Search Table: Contingency: Prospective data (chi-	$\sim$
<ul> <li>Recently used</li> <li>Transform, Normalize</li> <li>Transform concentrations (X).</li> <li>Normalize</li> <li>Prune rows</li> <li>Remove baseline and column math</li> <li>Transpose X and Y</li> <li>Fraction of total</li> <li>Stanalyses</li> <li>Column analyses</li> <li>Column analyses</li> <li>Column analyses</li> <li>Contingency table analyses</li> <li>Contingency table analyses</li> <li>Parts of whole analyses</li> <li>Multiple variable analyses</li> <li>Multiple variable analyses</li> <li>Generate curve</li> <li>Simulate data</li> </ul>	
	)
Help Cancel OK	

## 科研技能 单元课

### 卡方检验

#### 2 选择实验参数

Parameters: Chi-square (and Fisher's exact) test ×				
Main Calculations Options				
Effect sizes to report				
(1) Relative Risk Used for prospective and experimental studies				
(2) Difference between proportions (attributable risk) and NNT (3) Used for prospective and experimental studies				
(4) Odds ratio 2 Used for retrospective case-control studies				
Sensitivity, specificity and predictive values Used for diagnostic tests				
Method to compute the P value				
◯ Fisher's exact test	IN			
○ Yates' continuity corrected chi-square test				
Chi-square test 3				
O Chi-square test for trend	님			
Looking for the z test to compare proportions? Choose the chi-square test (with or without the Yates' correction). The chi-square and z tests are equivalent.				
(4)	È			
Learn Cancel OK				

①选择Main Calculation	
②选择Odds ratio	
③选择Chi-square test	
④选择OK	
(1) RR (Relative Risk	ҟѲӡҭ҅

(1) RR(Relative Risk ,相对危险度):暴露组研究对象的发病风险是非暴露组研究对象的多少倍

(2)AR(Attributable Risk , 归因危险度): 暴露组发病率与非暴露组发病率之差,它反映发病归因于暴露因 素的程度

(3) NNT(Number Needed to Treat ,需治疗人数): 可把抽象的率转变为 1 个具体的频数,使临床试验结 果转化为临床实践应用的指标,具有表达统计学意义及临床意义的双重作用,是一个衡量临床治疗效果、指导 临床决策的有用工具

(4) OR( Odds Ratio , 比值比): 在回顾性研究中可以评价暴露因素(阿司匹林)和疾病(心肌梗死)的 关联强度

## 科研技能 单元课

www.helixlife.cn/

### 卡方检验

#### 3 查看结果

1	Contingency	Α	В	С
1	Table Analyzed	Contingency: Prospective data (chi-square test)		
2				
3	P value and statistical significance			
4	Test	Chi-square		
5	Chi-square, df	25.01, 1		
6	Z	5.001		
7	P value	<0.0001		
8	P value summary	****		
9	One- or two-sided	Two-sided		
10	Statistically significant (P < 0.05)?	Yes		
11				
12	Effect size	Value	95% CI	
13	Odds ratio	1.832	1.439 to 2.338	
14	Reciprocal of odds ratio	0.5458	0.4277 to 0.6948	
15				
16	Methods used to compute CIs			
17	Odds ratio	Baptista-Pike		
18				
19	Data analyzed	Myocardial Infarction	No MI	Total
20	Placebo	189	10845	11034
21	Aspirin	104	10933	11037
22	Total	293	21778	22071
23				
24	Percentage of row total	Myocardial Infarction	No MI	
25	Placebo	1.71%	98.29%	
26	Aspirin	0.94%	99.06%	

## 科研技能 单元课

www.helixlife.cn/

【课程目录



解螺旋 | 陪伴医生科研成长

## 科研技能 单元课

相关分析

- 示例数据来源于Graphpad
- 每一行代表不同的日子。X柱表示臭氧水平。Y柱表示太阳辐射、风和温度
- 目标: 了解这三个天气变量与臭氧水平之间的关系

	Х	Group A	Group B	Group C
(1)	Ozone	Solar.R	Wind	Temp
	Х	Y	Y	Y
1	41	190	7.4	67
2	36	118	8.0	72
3	12	149	12.6	74
4	18	313	11.5	62
5			14.3	56
6	28		14.9	66
7	23	299	8.6	65
8	19	99	13.8	59
9	8	19	20.1	61
10		194	8.6	69
11	7		6.9	74
12	16	256	9.7	69
13	11	290	9.2	66
14	14	274	10.9	68
15	18	65	13.2	58
16	14	334	11.5	64
17	34	307	12.0	66
18	6	78	18.4	57
19	30	322	11.5	68
20	11	44	9.7	62
21	1	8	9.7	59
22	11	320	16.6	73
23	4	25	9.7	61
24	32	92	12.0	61
25		66	16.6	57
26		266	14.9	58
27			8.0	57

①录入数据(此处可以非完整数据,也就是说允许数据的缺少)
 ②点击分析按钮

#### 1 调出分析窗口

Clip	board Ar	nalysis	Change	Import Draw
of	* 🖻 🐁 🗹 🖾 🔂		E∎ ⊇∎ 2↓ - 💩 - 🌈 🖓	
Ē		vze 🏼 🏄 📑	<sup>₽</sup> #.# <sup>123</sup> <del>1.23</del> ()	× xml
	Х	Group A	Group B	Group C
	Ozone	Solar.R	Wind	Temp
	Х	Y	Y	Y
1	41	190	7.4	67
2	36	118	8.0	72
3	12	149	12.6	74
4	18	313	11.5	62
5			14.3	56
6	28		14.9	66
7	23	299	8.6	65
8	19	99	13.8	59
9	8	19	20.1	61
10		194	8.6	69
11	7		6.9	74
12	16	256	9.7	69

## 科研技能 单元课

相关分析

### 2 选择分析方法

Analyze Data	$\times$
Use: Built-in analysis V	
Which analyze which data sets?	
Analyze which data sets?	
Search Table: XY: Correlation	$\sim$
Recently used     A:Solar.R	
<ul> <li>Transform, Normalize</li> <li>Transform</li> <li>Transform concentrations (X)</li> <li>Normalize</li> <li>Prune rows</li> <li>Remove baseline and column math</li> <li>Transpose X and Y</li> <li>Fraction of total</li> <li>XY analyses</li> <li>Nonlinear regression (curve fit)</li> <li>Simple linear regression</li> <li>Simple logistic regression</li> <li>Fit spline/LOWESS</li> <li>Smooth, differentiate or integrate curve</li> <li>Area under curve</li> <li>Deming (Model II) linear regression</li> <li>Row statistics</li> <li>Correlation</li> <li>Interpolate a standard curve</li> <li>Counn analyses</li> <li>Grouped analyses</li> </ul>	
Contingency table analyses	
Survival analyses	
Deste of whole analysis      Desteict All     Desteict All     Desteict All	<b>A</b> II
	3)
Help Cancel	ОК

_						
P	arameters: Correlation	×				
	Compute correlation between which pairs of columns?					
	O Compute r for every pair of Y data sets (Correlation matrix).					
	When a value is missing or excluded, remove the entire row from the calculation					
4	Compute r for X vs. every Y data set:					
	🔨 [X] Ozone	J				
	O Compute r between two selected data sets:					
	🔨 [X] Ozone					
	🔨 [A] Solar.R					
	Ves. Compute Pearson correlation coefficients.     No. Compute nonparametric Spearman correlation.     Options     P value: One-tailed Two-tailed					
	Show this many significant digits (for everything except P values):	▲ ▼				
	P value style: GP: 0.1234 (ns), 0.0332 (*), 0.0021 (**) $\vee$ N = 6	*				
	Graphing					
	Create a heatmap of the correlation matrix.					
	Make these choices the default for future analyses					
	Learn Cancel OK					
1						

分析参数选择

(3)

- 选择相关分析
   勾选所有组别
   选择OK
   选择臭氧作为X(三个天气变量Y与臭氧 水平X之间的关系)
   根据是否符合高斯分布(即正态分布) 选择皮尔森系数或者斯皮尔曼系数
  - ⑥ 其它选项默认,选择OK

## 科研技能 单元课

www.helixlife.cn/

### 相关分析

#### 查看分析结果

Search ~			А	В	С
• Data Tables »	•	Correlation	Ozone vs.	Ozone vs.	Ozone vs.
XY: Correlation			Solar.R	Wind	Temp
🕀 New Data Table	4				
/ Info »	1	Pearson r			
(i) Project info 1	2	r	0.3483	-0.6015	0.6984
New Info	3	95% confidence interval	0.1732 to 0.5021	-0.7064 to -0.4709	0.5913 to 0.7812
Results »	4	R squared	0.1213	0.3619	0.4877
Correlation of XY: Correlation	5				
New Analysis	6	P value			
Graphs »	7	P (two-tailed)	0.0002	<0.0001	<0.0001
🕀 New Graph	8	P value summary	***	****	****
✓ Layouts »	9	Significant? (alpha = 0.05	Yes	Yes	Yes
New Layout	10				
	11	Number of XY Pairs	111	116	116

R平方,即相关系数的平方, 也就是决定系数,越接近1 表示相关性越强

## 科研技能 单元课

www.helixlife.cn/

【课程目录



解螺旋 | 陪伴医生科研成长

## 科研技能 单元课

### 线性回归分析

- 所有模型均将结果(Y)定义为一个或多项参数和一个独立变量(X)或几个独立变量的函数。目的是调整模型参数值,以找到最接近数据的直线或曲线 例如,对于线性回归,目的是找到斜率和截距的最佳拟合值,以使直线接近数据。 对标准化剂量-反应曲线进行非线性回归,目的是调整EC50的值(在最小和最大反应之间激发反应的浓度)和曲线斜率
- 线性回归示例:文献: Borkman M, Storlien LH, Pan DA, Jenkins AB, Chisholm DJ, Campbell LV. The relation between insulin sensitivity and the fatty-acid composition of skeletal-muscle phospholipids. N Engl J Med. 1993 Jan 28;328(4):238-44. doi: 10.1056/NEJM199301283280404. PMID: 8418404. Figure3a
  - 示例解读: X值是肌肉中某一种脂肪酸的百分比(来自活组织检查), Y值是胰岛素敏感性
  - 目的: 用线性回归来拟合数据中的直线
  - 在许多情况下,X是一个需要操作的变量,Y是一个需要测量的变量。在本例中这两个变量都被测量了,并且并不完全清楚哪个应该是X,哪个应该是Y



## 科研技能 单元课

www.helixlife.cn/

### 线性回归分析

#### 1 建立XY数据表

#### ① 选择XY表;②选择新创建数据;③X选择数字;④Y中每个点输入独立的Y值;⑤点击Create后,右图内键入数据

	New Data Table and Graph		>
		XY tables: Each point is defined by an X and Y coordinate	
	CREATE	X         A         B         Control         Treated           Minutes         Control         Treated         Treated         Treated	
(1)	XY Column	1 Titlo 2 Titlo 2 Titlo 2 Marco 2 Marc	
	Grouped	3 Title Learn more	
	Contingency	Data table:	
	Survival	Enter or import data into a new table     2	
	Parts of whole Multiple variables	Start with sample data to follow a tutorial	
	Nested	Options:	
		X: Numbers 3	
	OPEN	O Numbers with error values to plot horizontal error bars	
	Clone a Graph	Dates     Elarcot times	
		Y: O Enter and plot a single Y value for each point (4)	
		Enter and plot error values already calculated elsewhere	
		Enter: Mean, SD, N	
			5
	Prism Tips	Cancel	Create

_		Х	Group A
Table format:		%C20-22 POLYUNSATURATED FATTY ACIDS	INSULIN SENSITIVITY (MG/M <sup>2</sup> /MIN)
	×	X	Y
1	Title	17.9	250
2	Title	18.3	220
3	Title	18.3	145
4	Title	18.4	115
5	Title	18.4	230
6	Title	20.2	200
7	Title	20.3	330
8	Title	21.8	400
9	Title	21.9	370
10	Title	22.1	260
11	Title	23.1	270
12	Title	24.2	530
13	Title	24.4	375

## 科研技能 单元课

www.helixlife.cn/

### 线性回归分析

### 2 选择分析方法

ılyze Data					×
Use:	Built-in analysis		~	]	
/hich analysis?		Analyze	e which da	ata sets?	
earch		Table:	XY: Simp	ole linear regressi	on 🗸
Recently used	^		INSULIN S	ENSITIVITY (MG	/M <sup>2</sup> /MIN)
= Transform, Normalize					
Transform		(2)			
Transform concentrations	(X)				
Normalize					
Prune rows					
Remove baseline and colur	nn math				
Transpose X and Y					
Fraction of total		Wh	en vou an	alvze tables or gr	anhs with
XY analyses		mor	e than one	e data set, use thi	s space to
Nonlinear regression (curv	a fit)	s	elect whic	h data set(s) to a	nalyze.
Simple linear regression					
Simple logistic regression					
Smooth differentiate or in					
Area under curve	legrate curve				
Deming (Model II) linear re	aression				
Row statistics					
Correlation					
Interpolate a standard curv	/e				
E Column analyses					
Brouped analyses					
Contingency table analyse	s				
Survival analyses			Select All	Des	elect All
Darte of whole analyses	•		2010/01/11	000	0000111
					(3)
		н	elp	Cancel	ок

### 3 设置统计参数(可默认)

'a	rameters: Simple Linear Regression
	Interpolate
	Interpolate unknowns from standard curve
	Compare
	Test whether slopes and intercepts are significantly different
	Graphing options
	Show the 95% confidence bands $\checkmark$ of the best-fit line
	Residual plot
	Constrain
	Force the line to go through $X=$ 0 , $Y=$ 0
	Replicates
	Consider each replicate Y value as an individual point
	Only consider the mean Y value of each point
	I est departure from linearity with runs test
	95% confidence interval of Y when $X = 0$
	95% confidence interval of X when $Y = 0$
	Range Start regression line at: End regression line at:
	x = 17.9 $x = 24.4$
	Output
	Show this many significant digits (for everything except P values):
	P value style: (CP: 0.1234 (ns) 0.0332 (*) 0.0021 (**) 0 $\vee$ N = 6
l	
	Make these choices as default for future regressions
	More choices Learn Cancel OK

- 选择简单线性回归
   勾选所有组别
   由于Y方向的组别只有1组,因此默
   认已选上这一组,且不能修改
   若有多个组别,这里可以选择一到
   多个组别
   其它选项默认,选择OK
   通过标准曲线填补缺失数据
   检验斜率和截距是否显著不同
   添加某个CI值的线
   绘制残差图
- (5)可强迫拟合的直线经过某个点
- (6)当有重复实验时,这里可设置用平均值 或者是各个独立的实验值进行分析
- (7) 计算指定数据点的95%CI
- (8) 可设置在某段数据范围内进行回归分析
- (9) 显示回归分析中XY对应的值

## 科研技能 单元课

www.helixlife.cn/

### 线性回归分析

#### 4 查看分析结果

✓ Data Tables >>	>		Α
🖽 XY: Simple linear regression	1	Simple linear regression Tabular results	INSULIN SENSITIVITY
🕀 New Data Table		rabata resarts	(MG/M <sup>2</sup> /MIN)
✓ Info >>			
<ol> <li>Project info 1</li> </ol>	10		
New Info	11	95% Confidence Intervals	
✓ Results >>	, 12	Slope	16.75 to 57.67
📄 Simple linear regression of XY:.	. 13	Y-intercept	-912.9 to -60.18
New Analysis	14	X-intercept	3.562 to 15.97
✓ Graphs >>	15		
XY: Simple linear regression	16	Goodness of Fit	
New Graph	17	R squared	0.5929 (1)
v Lavouts	18	Sy.x	75.90
New Lavout	19		
	20	Is slope significantly non-zero?	
	21	F	16.02
	22	DFn, DFd	1, 11
	23	P value	0.0021 (2)
	24	Deviation from zero?	Significant
 Family >>	25		
XY: Simple linear regression	26	Equation	<b>Y = 37.21*X - 486.5</b> (3)
Simple linear regression	27		
XY: Simple linear regression	28	Data	
	29	Number of X values	13
	30	Maximum number of Y replicates	1
	31	Total number of values	13
	32	Number of missing values	0

#### (1) R squared,即相关系数的平方,也就是决定系

数,越接近1表示相关性越强

(2) *p*值

(3)回归方程:斜率<0表示是负相关

斜率>0表示正相关

## 科研技能 单元课

标记分析结果

线性回归分析



#### ① 点击Graphs下XY; ②选择散点图; ③点击OK; ④Y用文本框工具,将R平方和p值标记在图上



## 科研技能 单元课

### 非线性回归分析

● 所有模型均将结果(Y)定义为一个或多项参数和一个独立变量(X)或几个独立变量的函数。目的是调整模型参数值,以找到最接近数据的直线或曲线。 例如,对于线性回归,目的是找到斜率和截距的最佳拟合值,以使直线接近数据。 对标准化剂量-反应曲线进行非线性回归,目的是调整EC50的值(在最小和最大反应之间激发反应的浓度)和曲线斜率

#### ● 非线性回归示例:

示例解读:

- X列记录时间;每个时间点相对应三次数据,用于数据质量。有些单元格是空白的,以表示缺失的数据
- 目的: 拟合指数衰减曲线以确定衰减速率常数以及Y的起始值和平台值

Table	format:	Х	Group A				Group B	
>	KY (	Minutes	Control			Treated		
	X	Х	A:Y1	A:Y2	A:Y3	B:Y1	B:Y2	B:Y3
1	Title	1.0	8887	7366	9612	6532	7905	7907
2	Title	2.0	8329		8850	5352	5841	6277
3	Title	3.0	7907	8810	8669	5177	4082	3157
4	Title	4.0	7413	8481	6489	3608		4226
5	Title	5.0	7081	7178	5716	2559	3697	2816
6	Title	6.0	6249	6492		1671	3053	2891
7	Title	8.0	5442	6172	6409	2264	1658	1879
8	Title	10.0	4020	3758	4138	1905	1302	1406
9	Title	14.0	4559	3146	2547	2994	1338	739
10	Title	20.0	3033	1587	2754	1444		760
11	Title	25.0	2105	1707	2152	281	484	765
12	Title	30.0	1005	2156	1185	1103	1517	833
13	Title	50.0	820	1513	1591	1918	1128	1293

## 科研技能 单元课

非线性回归分析

🚺 建立XY数据表,复制数据到表中





## 科研技能 单元课

非线性回归分析



Para	meters: Nonlinear Regression				
-	- (1)				
Mo	del Method Compare Constrain Initial values Range	Output	Confidence	Diagnostic	s Flag
Ch	oose an equation				
	Standard curves to interpolate		-	Ne	N -
	Dose-response - Stimulation				
	Dose-response - Inhibition			Detai	le
	Bose-response - Special, X is concentration			Detai	
	Dose-response - Special, X is log(concentration	)			
	Binding - Saturation				
	Binding - Competitive				
	Binding - Kinetics				
	Enzyme kinetics - Inhibition				
	Enzyme kinetics - Velocity as a function of subs	trate			
	Exponential				
	Distance decay				
	Plateau followed by one phase decay				
	Three phase decay				
	Plateau followed by one phase association				
	Two phase association				
	Exponential growth equation				
	Ines				
	Polynomial		~		
			_		
	If you have subtracted off the nonspecific signal, constra	n Plateau	to a constar	nt value of 0.	0
	One phase decay				
	Analytical derivatives		2 Loarn ab	out this oqua	tion
1			Eeann ab	out this eque	
In	terpolate				
[	Interpolate unknowns from standard curve. Confidence	interval:	None $\sim$		3
		Lea	rn	Cancel	ОК
				(	



Х

-	Nonlin fit	A	В
	Table of results	Control	Treated
1	One phase decay		
2	Best-fit values		
3	Y0 (1)	9992	9593
4	Plateau (2)	987.0	1154
5	к (3)	0.08927	0.3042
6	Half Life (4)	7.765	2.278
7	Tau (5)	11.20	3.287
8	Span (6)	9005	8440
9	95% CI (profile likelihood	)	
10	YO	9301 to 10738	8419 to 10967
11	Plateau	201.0 to 1638	813.2 to 1480
12	К	0.06862 to 0.1134	0.2387 to 0.3862
13	Half Life	6.110 to 10.10	1.795 to 2.903
14	Tau	8.815 to 14.57	2.589 to 4.189
15	Goodness of Fit		
16	Degrees of Freedom	34	34
17	R squared	0.9401	0.9196
18	Sum of Squares	16891590	12555343
19	Sy.x	704.8	607.7
20	Constraints		
21	К	K > 0	K > 0
22			
23	Number of points		
24	# of X values	39	39
25	# Y values analyzed	37	37

① 选择model

② 选择One phase decay

3 OK

(1) Y0是X(时间)为0时的Y值。 与Y用相同的单位表示

(2) Plateau是无穷次时的Y值,

与Y用相同的单位表示

(3) K是速率常数,用X轴时间单 位的倒数表示

(4) 半衰期在X轴的时间单位内, 它被计算为ln(2)/K

(5)Tau是时间常数,用与X轴相 同的单位表示

(6) Span是Y0和Plateau之间的 差值,用与Y值相同的单位表示

# 科研技能单元课

### 生存曲线差异分析

- 仅用于生存分析,由观察对象的生存时间和生存状态构成,1表示终点事件,0表示删失数据,注意分组数据的以"缺失数据"的形式进行表示
- 示例: 文献Han J, Wang JZ, Yang X, Yu H, Zhou R, Lu HC, Yuan WB, Lu JC, Zhou ZJ, Lu Q, Wei JF, Yang H. METTL3 promote tumor proliferation of bladder cancer by accelerating pri-miR221/222 maturation in m6A-dependent manner. Mol Cancer. 2019 Jun 22;18(1):110. doi: 10.1186/s12943-019-1036-9. PMID: 31228940; PMCID: PMC6588935. Fig.1f
- 示例解读:
  - > 图中实验是180例膀胱癌根据mettle3表达量的多少分为low和 high两组,对存活时间进行记录
  - ▶ 图中X轴作为生存时间,Y轴表示存活率



Low expression	Survival time (months)	High expression	Survival time (months)
1	17	1	10
1	38	1	3
1	26	1	10
1	34	1	6
1	36	1	27
1	25	1	7
1	20	1	18
1	14	1	26
1	27	1	18
1	20	1	14
1	15	1	16
1	29	1	22
1	18	1	15
1	20	1	25
1	42	1	26
1	26	1	32
1	14	1	40
1	36	1	34

• 模拟数据

## 科研技能 单元课

### 生存曲线差异分析

1	Survival Curve comparison	А	В
	Curve companson		
4	Comparison of Survival Curves		
2	companson of Survival Curves		
3	Log-rank (Mantel-Cox) test		
4	Chi square	7.373	
5	df	1	
6	P value	0.0066	
7	P value summary	**	
8	Are the survival curves sig different?	Yes	
9			
10	Gehan-Breslow-Wilcoxon test		
11	Chi square	6.420	
12	df	1	
13	P value	0.0113	
14	P value summary	*	
15	Are the survival curves sig different?	Yes	
16			
17	Median survival		
18	Low expression	Undefined	
19	High expression	43.00	
20			
21	Hazard Ratio (Mantel-Haenszel)	A/B	B/A
22	Ratio (and its reciprocal)	0.5589	1.789
23	95% CI of ratio	0.3672 to 0.8506	1.176 to 2.723
24			
25	Hazard Ratio (logrank)	A/B	B/A
26	Ratio (and its reciprocal)	0.5619	1.780
27	95% Cl of ratio	0.3705 to 0.8523	1.173 to 2.699

GraphPad Prism在作生存曲线时,便自动对数据进行分析;对于生存曲线而言,在使用的时候更重要的 是对数据的录入,相关内容的介绍在作图教程中有详细解答

- 给出了两种检验方法的结果:
- Log-rank检验(常用):检验对不同生存时间点均取权重为"1",即不考虑各观察时间点开始时存活的人数对统计模型的影响。相对重视远期效应
- Gehan-Breslow-Wilcoxon检验:在Log-rank检验的基础上增加了权重,按各个时间点存活的人数加权 即开始存活人数相对多的时间点死亡情况的变化对整个模型的贡献较大,而开始存活人数相对少的时点 死亡情况的变化对整个模型的贡献较小。对近期效应更加敏感。

(这里不理解不要紧,只要记住结论即可)

- 两个方法的结论相同: 可认为近期和远期都有差别(或无差别)
- 若Log-rank检验 p < 0.05而Gehan-Breslow-Wilcoxon检验 p > 0.05,可认为远期有差别而近期无差别,反

之,近期有差别而远期无差别

## 科研技能 单元课

小结: GraphPad Prism可简单便捷地实现科研数据处理

U检验 T检验 卡方检验 方差分析	皮尔森相关分析 斯皮尔曼方差分析 典型相关分析	线性回归 逻辑回归 有序逻辑回归 多项逻辑回归
假设检验	相关分析	回归分析



更多SCI相关证	果程:
《 <u>临床SCI零基础</u>	出撰写入门课程》
《 <u>TCGA数据挖掘</u>	主法演示》
《单细胞转录组	数据分析实战》
更多精彩课程,	将陆续推出









科研技能 单元课 03 增刊

针对医生做基础科研过程中需要的特定科研技能,解螺旋制作了一系列
 专项学习突破的图文或视频教程,谓之"单元课",此为第3期的增刊。



