

无统计，不科研

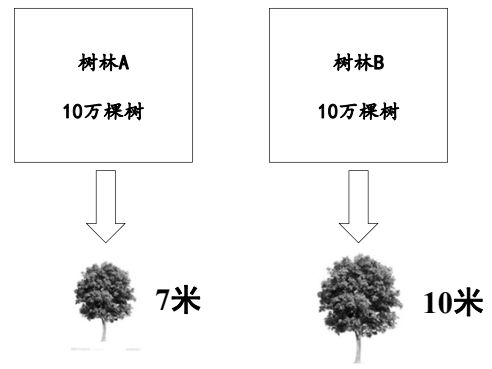
——最常用统计方法介绍

吕军 副主任

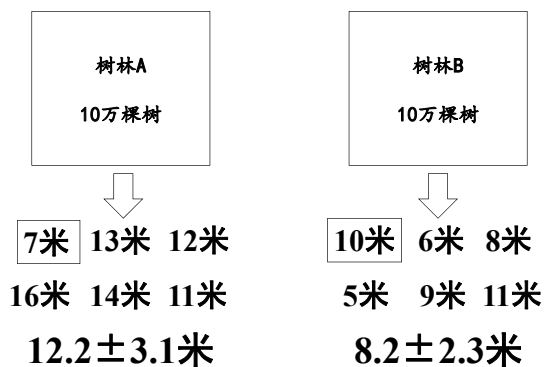
西安交通大学第一附属医院 临床研究中心

2015年10月18日

统计的意义



统计的意义



统计的意义

科学研究的目的是要阐明总体的特征和规律，但实际工作中多数情况下没有必要或不可能对总体中的每一个个体进行观测。

科学的办法是进行抽样研究，即从总体中抽取一部分有代表性的个体，其观测值就构成样本。

对样本做深入研究，利用获得的样本信息进行统计推断，阐明总体的特征和规律。



统计书

厚重难懂



可解决90%科研问题的三种统计方法

1. t 检验：两组数据对比

计量
资料

2. 方差分析：三组及以上数据对比

3. 卡方检验：两组及以上数据对比

计数
资料

一、 t 检验

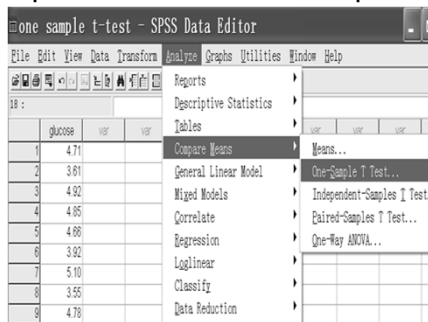
- 单样本 t 检验
- 配对 t 检验
- 两样本 t 检验

1. 单样本 t 检验

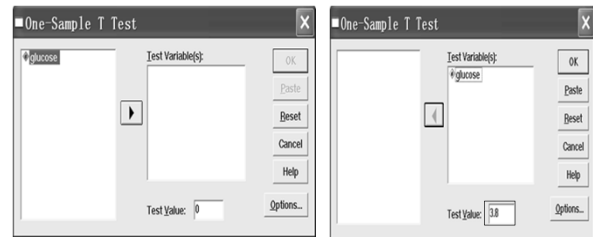
- 例1 为研究口服避孕药是否会增高血糖测量值，某研究人员随机抽取9名一直口服避孕药至少6个月的妇女，测量她们的空腹血糖值，结果如下所示。已知不服用避孕药妇女的空腹血糖值均值为3.8 mmol/L。根据该样本提供的信息，是否有证据表明口服避孕药可以提高空腹血糖值？

4.71 3.61 4.92 4.85 4.66
3.92 5.10 3.55 4.78

Compare Means → One Sample T Test



One Sample t Test对话框



结果解释

- 单样本 t 检验结果显示：（统计方法）
- 服用避孕药与不服用避孕药妇女空腹血糖值的差别有统计学意义（ $t=3.31$ ， $P=0.011$ ）（统计结论）
- 口服避孕药可以增加血糖水平，平均增加0.66（95% CI：0.20~1.11）mmol/L（专业结论）

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
blood glucose levels (mmol/L)	9	4.4556	.59374	.19791

One-Sample Test					
Test Value = 3.8					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower Upper
blood glucose levels (mmol/L)	3.312	8	.011	.6556	.1992 1.1119

2. 配对 t 检验（paired t -test）

配对设计：

- 同一受试对象处理前后
- 同一受试对象分别接受两种不同的处理
- 两个同质受试对象分别接受两种不同的处理

(1) 同一受试对象处理前后

- 例2 用家兔10只实验某批注射液对体温的影响。体温在注射前每隔1小时测定1次，连续2次；在注射后1小时和2小时分别各测定1次，得到每只家兔注射前后体温的均数。问该批注射液是否能引起家兔体温升高？

表1 某批注射液对家兔体温（℃）的影响

家兔编号	1	2	3	4	5	6	7	8	9	10
注射前体温	37.8	38.2	38.0	37.6	37.9	38.1	38.2	37.5	38.5	37.9
注射后体温	37.9	39.0	38.9	38.4	37.9	39.0	39.5	38.6	38.8	39.0

(2) 同一受试对象分别接受两种不同的处理

- 取20份样品，每份一分为二，分别用容量法和仪器分析测定。试判断容量法的测量结果是否低于仪器分析结果？
- 某医院用A、B两种血红蛋白测定仪器检测了16名健康男青年的血红蛋白含量（g/L），问两种血红蛋白测定仪器的检测结果是否有差别？
- 某研究者比较耳垂血和手指血的白细胞数，调查了12名成年人，同时采取耳垂血和手指血，试比较两者的白细胞数有无不同？

- 有12名接种卡介苗的儿童，8周后用两批不同的结核菌素，一批是标准结核菌素，一批是新制结核菌素，分别注射在儿童的前臂，两种结核菌素的皮肤浸润反应直径（mm）如表所示。问两种结核菌素的反应性有无差别？

表3 12名儿童分别用两种结核菌素皮肤浸润反应结果(mm)

编号	标准品	新制品	差值
1	12.0	10.0	2.0
2	14.5	10.0	4.5
3	15.5	12.5	3.0
4	12.0	13.0	-1.0
5	13.0	10.0	3.0
6	12.0	5.5	6.5
7	10.5	8.5	2.0
8	7.5	6.5	1.0
9	9.0	5.5	3.5
10	15.0	8.0	7.0
11	13.0	6.5	6.5
12	10.5	9.5	1.0

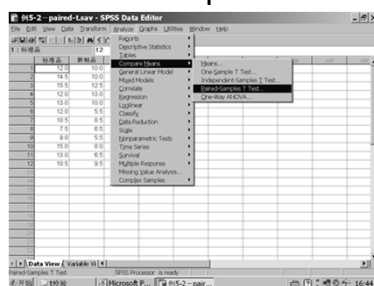
(3) 两个同质受试对象分别接受两种不同的处理

- 例3 为研究某种抗癌新药对小白鼠移植性肉瘤S₁₈₀的抑瘤效果，将20只小白鼠按性别、体重、窝别配成对子。每对中随机抽取1只服用抗癌新药“扶正消瘤汤”，另1只作为阴性对照，服用生理盐水，观察其对小白鼠移植性肉瘤S₁₈₀的抑瘤效果。经过一定时间，测得小白鼠瘤重如表所示。问小白鼠服用抗癌新药和生理盐水后平均瘤重有无不同？

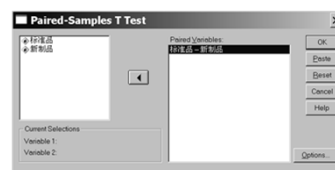
表2 抗癌新药“扶正消瘤汤”对小白鼠移植性肉瘤S₁₈₀的抑瘤效果（瘤重/g）

家兔编号	1	2	3	4	5	6	7	8	9	10
对照组	0.80	0.74	0.31	0.48	0.76	0.65	0.72	0.68	0.39	0.53
新药组	0.47	0.42	0.38	0.44	0.38	0.43	0.39	0.71	0.32	0.41

Analyze → Compare Means → Paired-Samples T test



将变量标准品和新制品成对选入 Paired Variables 栏；单击OK



- 配对 t 检验结果显示：新制品结核菌素皮肤浸润反应的直径均数低于标准品（ $t = 4.52$, $P = 0.001$ ），均数差值为 3.25（95% CI: 1.67~4.83）mm。

Paired Samples Statistics				
	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 标准品	12.042	12	2.3975	.6921
新制品	8.792	12	2.5067	.7242

Paired Samples Test							
Paired Differences							
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df
				Lower	Upper		
Pair 1 标准品 - 新制品	3.2500	2.4909	.7191	1.6674	4.8326	4.520	11
							Sig. (2-tailed)
							.001

3. 独立样本 t 检验

- 分别从两个研究总体中随机抽取样本，然后比较两样本某指标的均数；
- 随机抽取实验对象，将其随机分成两组，分别接受两种不同处理，然后比较两组的平均效应。

（1）分别从两个研究总体中随机抽取样本，然后比较两样本某指标的均数

- 例4 从两个品种的家兔分别测定其停药 18 小时后正常血糖值，问这两个品种家兔的正常血糖值有无差别？

表4 两个不同品种家兔的正常血糖值

家兔品种	正常血糖值 (mg/100ml)									
大耳白	57	120	101	137	119	117	104	73	53	68
青紫蓝	89	36	82	50	39	32	57	82	96	31

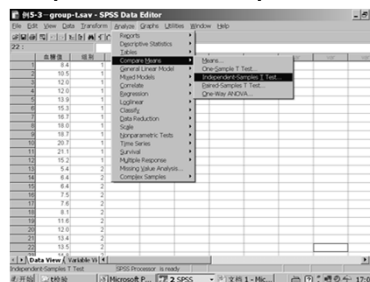
（2）随机抽取实验对象，将其随机分成两组，分别接受两种不同处理，然后比较两组的平均效应。

- 例5 25例糖尿病患者随机分成两组，甲组单纯用药物治疗，乙组采用药物治疗合并饮食疗法，两个月后再次测空腹血糖，问两组患者血糖值是否相同？

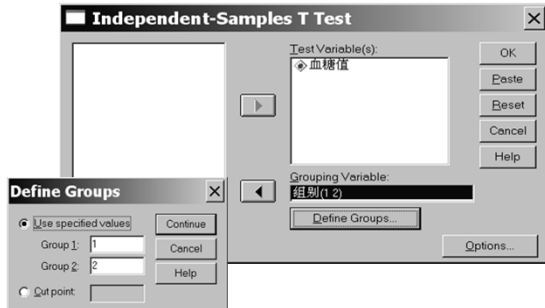
表5 25例糖尿病患者两种疗法后两个月血糖值 (mmol/L)

甲组血糖值 X_1	乙组血糖值 X_2
8.4	5.4
10.5	6.4
...	...
21.1	14.8
15.2	15.6
	18.7

Analyze → Compare Means → Independent-Samples T test



将变量血糖值选入Test栏；将组别选入grouping栏，单击Define Groups按钮，定义两个组的变量值，本例为1和2，单击Continue按钮；单击OK按钮。



- 经方差齐性检验， $F=0.28$ ， $P=0.60$ 。按 $\alpha=0.10$ 水准，可认为两总体方差齐性（如不齐用第2行数据）。
- 两样本 t 检验结果显示：单纯药物治疗的血糖值均数高于药物合并饮食治疗（ $t=2.64$ ， $P=0.015$ ），均数差值为4.36（95% CI: 0.94~7.78）mmol/L。

Group Statistics				
血糖值	组别	N	Mean	Std. Deviation
	甲组	12	15.208	4.0216
	乙组	13	10.846	4.2215

Independent Samples Test									
		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference
血糖值	Equal variances assumed	.283	.600	2.640	23	.015	4.3622	1.6522	.9444 7.7800
	Equal variances not assumed			2.646	22.972	.014	4.3622	1.6488	.9511 7.7733

4. t 检验的应用条件

- n 较小时（如 n_1 或 $n_2 < 60$ ），理论上要求样本取自正态总体。
- 两小样本均数比较时，要求两总体方差相等。

二、方差分析

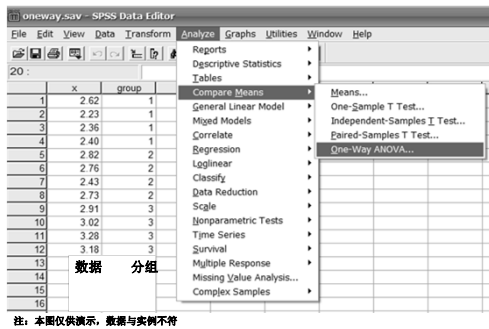
1. 完全随机设计资料的方差分析 *One-Way ANOVA*

例6 某高原研究组将籍贯相同、年龄相同、身高体重接近的30名新战士随机分为三组，甲组为对照组，按常规训练，乙组为锻炼组，每天除常规训练外，接受中速长跑与健身操锻炼，丙组为药物组，除常规训练外，服用抗疲劳药物，一月后测定第一秒用力肺活量(L)，结果见表。试比较三组第一秒用力肺活量有无差别。

表6 三组战士的第一秒用力肺活量(L)

对照组 (A)	锻炼组 (B)	药物组 (C)	合计
3.25	3.66	3.44	
3.32	3.64	3.62	
3.29	3.48	3.48	
3.34	3.64	3.36	
3.16	3.48	3.52	
3.64	3.20	3.60	
3.60	3.62	3.32	
3.28	3.56	3.44	
3.52	3.44	3.16	
3.26	3.82	3.28	

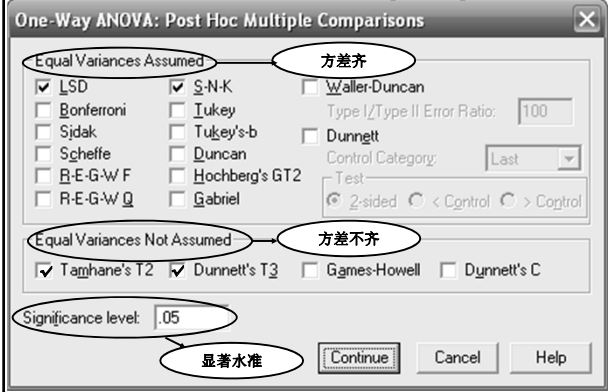
➤ 数据格式 n行2列 (指标变量、分组变量)



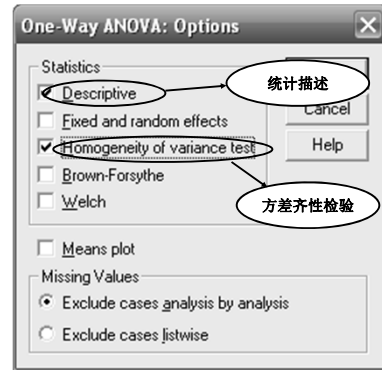
One-Way ANOVA 对话框



Post Hoc Multiple Comparisons 对话框



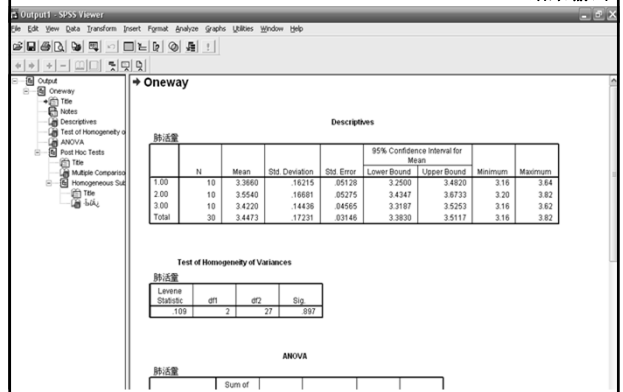
Option对话框



点击“OK”，运行结果



➤ 结果输出



基本统计描述

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1.00	10	3.3660	.16215	.05128	3.2500	3.4820	3.16	3.64
2.00	10	3.5540	.16681	.05275	3.4347	3.6733	3.20	3.82
3.00	10	3.4220	.14436	.04565	3.3187	3.5253	3.16	3.62
Total	30	3.4473	.17231	.03146	3.3830	3.5117	3.16	3.82

方差齐性检验

Levene Statistic of Homogeneity of Variances			
Levene Statistic	df1	df2	Sig.
.109	2	27	.897

方差分析表

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.186	2	.093	3.729	.037
Within Groups	.675	27	.025		
Total	.861	29			

多重比较

Post Hoc Tests

S-NK (齐)
Tamhane's T2 (不齐)
Dunnnett's T3 (不齐)

Dependent Variable: 肺活量

LSD	() 分组	(j) 分组	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
	1.00	2.00	-.18800*	.07069	.013	-.3330	-.0430
		3.00	-.05600	.07069	.435	-.2010	.0890
	2.00	1.00	.18800*	.07069	.013	.0430	.3330
		3.00	.13200	.07069	.073	-.0130	.2770
	3.00	1.00	.05600	.07069	.435	-.0890	.2010
		2.00	-.13200	.07069	.073	-.2770	.0130

*. The mean difference is significant at the .05 level.

三、 χ^2 检验

- 四格表资料的卡方检验
- 配对四格表资料的卡方检验
- 行×列表资料的卡方检验
 - 多个样本率比较的 χ^2 检验
 - 两个或多个样本构成比的比较

1. 四格表资料的卡方检验

- 例7 某医师欲比较胞磷胆碱与神经节苷酯治疗脑血管疾病的疗效，将78例脑血管病患者随机分为两组。结果如下：胞磷胆碱组52例，有效46例，有效率为88.5%；神经节苷酯组26例，有效18例，有效率为69.2%。问两种药物治疗脑血管疾病的有效率有无差别？

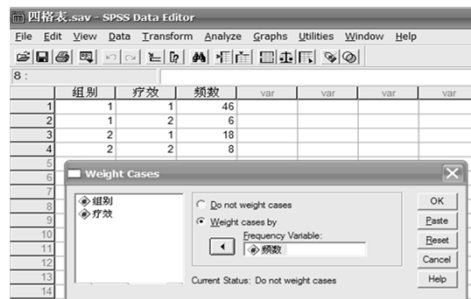
注：四个表中某数据期望频数小于5，需做连续性校正。

总例数小于40，需用Fisher确切概率检验。

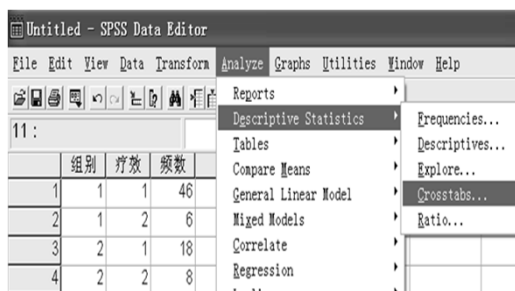
表7 两种药物治疗脑血管疾病有效率(%) 的比较

组别	有效	无效	合计	有效率
胞磷胆碱组	46	6	52	88.46
神经节苷酯组	18	8(4.67)	26	69.23
合计	64	14	78	82.05

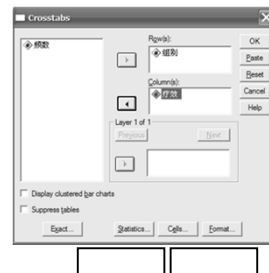
用变量“频数”加权：从菜单选择Data → Weight cases，打开Weight cases对话框，选择Weight cases by，将“频数”选入Frequency栏，单击OK 按钮



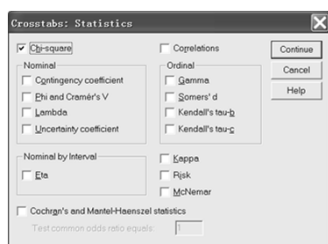
Analyze→Descriptive Statistics→Crosstabs



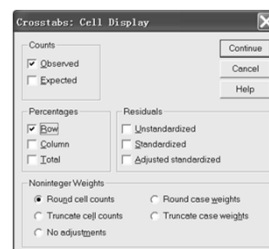
将行变量“分组”选入Row(s)栏，将列变量“疗效”选入Column栏，分别单击Statistics，Cell (格子)



选择 Chi-square



选择显示行百分数 Row



用四格表资料 χ^2 检验的校正公式计算得， $\chi^2 = 3.14$ ， $P = 0.076$ 。按 $\alpha = 0.05$ 水准，两种药物治疗脑血管疾病有效率的差别无统计学意义。

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.353 ^b	1	.037		
Continuity Correction ^a	3.145	1	.076		
Likelihood Ratio	4.126	1	.042		
Fisher's Exact Test				.058	.041
Linear-by-Linear Association	4.297	1	.038		
N of Valid Cases	78				

a. Computed only for a 2x2 table

b. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 4.67.

2. 配对四格表资料的卡方检验

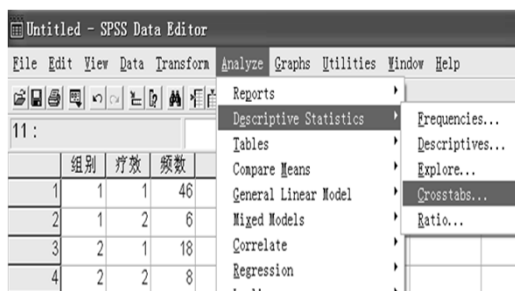
- 例8 某实验室分别用乳胶凝集法和免疫荧光法对58名可疑系统红斑狼疮患者血清中抗核抗体进行测定，结果见表。问两种方法的检测结果有无差别？

免疫：23/58 = 0.397
乳胶：13/58 = 0.224

表14 两种方法的检测结果

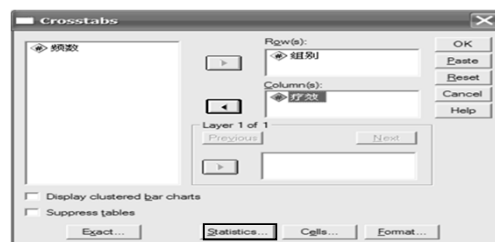
免疫荧光法	乳胶凝集法		合计
	+	-	
+	11 (a)	12 (b)	23
-	2 (c)	33 (d)	35
合计	13	45	58

Analyze→Descriptive Statistics→Crosstabs



注：本图仅供演示，数据与实际不符

将行变量“分组”选入Row(s)栏，
将列变量“疗效”选入Column(s)栏，
单击Statistics



注：本图仅供演示，数据与实际不符

配对四格表资料的 χ^2 检验，如在Crosstabs过程中进行，不显示配对 χ^2 检验的 χ^2 值，只显示四格表资料 χ^2 检验的 χ^2 值。



经配对四格表资料的 χ^2 检验， $P = 0.013$ 。按 $\alpha = 0.05$ 水准，两种方法的检测结果有差别，即免疫荧光法的检出率高于乳胶凝集法
McNemar：麦克尼马尔检验

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		.013 ^a
N of Valid Cases	58	

a. Binomial distribution used.

3. 行×列表资料的卡方检验 (多个样本率比较的 χ^2 检验)

- 例9: 某医师研究物理疗法、药物治疗和外用膏药三种疗法治疗周围性面神经麻痹的疗效。问三种疗法的有效率有无差别?

表8 三种疗法有效率的比较

疗法	有效	无效	合计	有效率(%)
物理疗法组	199	7	206	96.60
药物治疗组	164	18	182	90.11
外用膏药组	118	26	144	81.94
合计	481	51	532	90.41

SPSS数据

	疗法	疗效	freq
1	1	1	199
2	1	2	7
3	2	1	164
4	2	2	18
5	3	1	118
6	3	2	26

分组·疗效 Crosstabulation

分组	疗效	Total	
		有效	无效
物理疗法	Count	199	7
	% within 分组	96.6%	3.4%
药物治疗	Count	164	18
	% within 分组	90.1%	9.9%
外用膏	Count	118	26
	% within 分组	81.9%	18.1%
Total	Count	481	51
	% within 分组	90.4%	9.6%

结果解释

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	21.038 ^a	2	.000
Likelihood Ratio	21.559	2	.000
Linear-by-Linear Association	20.903	1	.000
N of Valid Cases	532		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.80.

- 经 χ^2 检验, $\chi^2=21.04$, $P=0.000$ 。按 $\alpha=0.05$ 水准, 可以认为三种疗法治疗周围性面神经麻痹的有效率有差别。

行×列表资料的卡方检验 (两个或多个样本构成比的比较)

- 例10 某医师在研究血管紧张素I转化酶(ACE)基因I/D多态与II型糖尿病肾病(DN)的关系时, 将249例II型糖尿病患者按有无糖尿病肾病分为两组, 资料见表6-9。问两组II型糖尿病患者的ACE基因型分布有无差别?

表9 DN组与无DN组2型糖尿病患者ACE基因型分布的比较

组别	DD	ID	II	合计
DN组	42 (37.8)	48 (43.3)	21 (18.9)	111
无DN组	30 (21.7)	72 (52.2)	36 (26.1)	138
合计	72 (28.9)	120 (48.2)	57 (22.9)	249

结果解释

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	7.913 ^a	2	.019
Likelihood Ratio	7.903	2	.019
Linear-by-Linear Association	6.447	1	.011
N of Valid Cases	249		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 25.41.

- 经 χ^2 检验, 可认为DN组与无DN组的II型糖尿病患者的ACE基因型分布不同 ($\chi^2=7.91$, $P=0.019$)。

卡方检验SPSS操作太繁琐, 怎么办?



别急, 有好东西!



注：

本课件所用素材的软件版本及软件
语言与现用软件有一定差别（大同小
异），请勿完全对号入座。

谢谢聆听！

西安交大一附院

临床研究中心

029-85323614

xjtuyfy_crc@163.com