

实验三：卡方检验与拟合优度检验

实验目的：了解做单侧检验的方法，了解方差分析显著性与功效的区别与联系，掌握用 SPSS 做卡方检验与拟合优度检验的基本步骤，能正确解读分析结果，规范书写检验报告。继续熟练数据的格式变换，掌握变量参数的设置对报告输出可读性的影响。

关键：选用合适的检验方法，正确解读分析结果

1. 均值检验的一些拓展（此内容选作，不要求写实验报告）

1.1 单侧检验

拓展题：某医生测量了 36 名从事铅作业男性工人的血红蛋白含量，算得其均数为 130.83g/L，标准差为 25.74g/L。问从事铅作业工人的血红蛋白是否小于正常成年男性平均值 140g/L ($\alpha=0.05$)？

方法一：分析→比较平均值→单样本 t 检验，观察结果中的 t 分布界值表，同一自由度，同一 t 值下，单侧检验 p 值减半；

方法二：分析→比较平均值→单样本 t 检验，将结果中的 t 值，与 $t_{35,0.05}$ 的理论值对比；

方法三：分析→比较平均值→单样本 t 检验，将理论均值与 90%置信区间的上限比较；

1.2 方差分析 差异显著性 (p) 与 功效 (η^2)

$$\eta^2 = SS_{\text{between}} / SS_{\text{total}}$$

$$\eta^2 = F * (k - 1) / [F * (k - 1) + (n - k)]$$

问题 1：比较对照组和处理组中 GTP 与 GOP 是否有差异，哪一个差异更显著？

问题 2：比较对照组与处理组中 GTP 与 GOP，哪一个差异更大？

问题 3：GTP 与 GOP 哪一个对处理更敏感？

组别	样本含量	对照	处理
GTP	5	16±3	20±3
GOP	200	16±3	18±3

2. 卡方检验

2.1 2×2 列联表的独立性检验

例题 7.3 某药有两种给药方式，口服给药和注射给药，下表是不同给药方式与给药结果的统计结果，问口服给药与注射给药的效果是否有差异。

给药方式	有效 (A)	无效 (\bar{A})	总数	有效率
口服 (B)	58	40	98	59.2%
注射 (\bar{B})	64	31	95	67.4%
总数	122	71	193	

1) 将表中的数据输入 SPSS，由于只有四个数字，可以直接输入，注意输入后将会有三列，一列是给药方式，一列是是否有效，最后一列为频数。

2) 检验步骤：

非常重要但易遗忘步骤：数据 → 加权个案 → 加权个案：频数变量：频数 → 确定

分析 → 统计描述 → 交叉表格：行：给药方式；列：效果 → 点击 statistics，勾选卡方 → 继续 → 点击单元格，勾选观察值与期望值 → 继续 → 确定

3) 结果解读

当 $n \geq 40$ 且所有 $T \geq 5$ 时，查阅 Pearson Chi-Square（皮尔逊卡方）

当 $n \geq 40$ 但有 $1 \leq T < 5$ 时，（或 $df=1$ 时）查阅 continuity correction（连续校正）

当 $n < 40$ 或 $T < 1$ （或 $p \approx \alpha$ 时），查阅 Fisher's Exact Test（Fisher 精确检验）

4) 书写检验报告

$H_0: \pi_1 = \pi_2$ ，不能认为两种给药方式效果有显著差异

$H_1: \pi_1 \neq \pi_2$ ，可以认为两种给药方式效果有显著差异

$\alpha=0.05$

输出结果：

个案处理摘要						
	个案					
	有效		缺失		总计	
	N	百分比	N	百分比	N	百分比
给药方式 * 效果	193	100.0%	0	0.0%	193	100.0%

给药方式 * 效果 交叉表					
		效果		总计	
		有效	无效		
给药方式	口服	计数	58	40	98
		期望计数	61.9	36.1	98.0
	注射	计数	64	31	95
		期望计数	60.1	34.9	95.0
总计	计数	122	71	193	
	期望计数	122.0	71.0	193.0	

卡方检验					
	值	自由度	渐进显著性 (双侧)	精确显著性 (双侧)	精确显著性 (单侧)
皮尔逊卡方	1.390 ^a	1	.238		
连续性修正 ^b	1.060	1	.303		
似然比	1.392	1	.238		
费希尔精确检验				.296	.152
线性关联	1.382	1	.240		
有效个案数	193				
a. 0 个单元格 (.0%) 的期望计数小于 5。最小期望计数为 34.95。					
b. 仅针对 2x2 表进行计算					

检验结论： $n \geq 40$ 且所有 $T \geq 5$ ，需要查阅皮尔逊卡方。由于 $p=0.238 > 0.05$ ，拒绝 H_1 ，接受 H_0 ，因此不能认为两种给药方式有显著差异。

2.2 配对四格表的卡方检验

现有 198 份痰标本，每份标本分别用 A、B 培养基培养结核菌，结果见下表。问 A、B 两种培养基的阳性培养率是否不等？

A 培养基	B 培养基		合计
	+	-	
+	48	24	72
-	20	106	126
合计	68	130	198

1) 将表中的数据输入 SPSS，由于只有四个数字，可以直接输入，注意输入后将会有三列，一列是 A 培养基，一列是 B 培养基，最后一列为频数。

2) 检验步骤：

数据 → 加权个案 → 加权个案：频数变量：频数 → 确定

分析 → 统计描述 → 交叉表格：行：A 培养基；列：B 培养基 → 点击 statistics，勾选卡方和 McNemar → 继续 → 点击单元格，勾选观察值与期望值 → 继续 → 确定

3) 结果解读

注意：输入的数据中除频数外的两列可以是字符串的数值类型，但在卡方的输出报告中会根据值进行自动排序，如果一定要按照指定顺序显示，可将值定义为 1, 2, 3 等有序数字，然后添加值标签，可确保显示的顺序按照预期排列。

4) 书写检验报告

$H_0: \pi_1 = \pi_2$ ，不能认为两种培养基的阳性培养率有显著差异

$H_1: \pi_1 \neq \pi_2$ ，可以认为两种培养基的阳性培养率有显著差异

$\alpha=0.05$

输出结果：

个案处理摘要						
	个案					
	有效		缺失		总计	
	N	百分比	N	百分比	N	百分比
A培养基 * B培养基	198	100.0%	0	0.0%	198	100.0%

A培养基 * B培养基 交叉表					
			B培养基		总计
			+	-	
A培养基	+	计数	48	24	72
		期望计数	24.7	47.3	72.0
	-	计数	20	106	126
		期望计数	43.3	82.7	126.0
总计		计数	68	130	198
		期望计数	68.0	130.0	198.0

卡方检验					
	值	自由度	渐进显著性 (双侧)	精确显著性 (双侧)	精确显著性 (单侧)
皮尔逊卡方	52.425 ^a	1	.000		
连续性修正 ^b	50.196	1	.000		
似然比	52.819	1	.000		
费希尔精确检验				.000	.000
线性关联	52.160	1	.000		
麦克尼马尔检验				.652 ^c	
有效个案数	198				

a. 0 个单元格 (.0%) 的期望计数小于 5。最小期望计数为 24.73。

b. 仅针对 2x2 表进行计算

c. 使用了二项分布。

检验结论：做 McNemar 检验，由于 $p=0.652>0.05$ ，拒绝 H_1 ，接受 H_0 ，即不能认为两种培养基的阳性培养率有显著差异。

2.3 行(R) X 列(C) 表资料的卡方检验及事后多重检验 (此内容选作, 不要求写实验报告)

某研究人员收集了亚洲、欧洲和北美洲人的 A、B、AB、O 血型资料，结果如下表，问不同地区人群 ABO 血型分类构成比是否不同。

地区	A	B	AB	O	合计
亚洲	321	369	95	295	1080
欧洲	258	43	22	194	517
北美洲	408	106	37	444	995
合计	987	518	154	933	2592

1) 将表中的数据输入 SPSS，调整最终的格式为纵向结构

2) 检验步骤：

数据 → 加权个案 → 加权个案：频数变量：频数 → 确定

分析 → 统计描述 → 交叉表格：行：地区；列：血型 → 点击 statistics，勾选卡方 → 继续 → 点击单元格，勾选观察值与期望值 → 继续 → 确定

3) 结果解读

行列表中的 $n \geq 40$ ，且各格 $T \geq 1$ ，并且 $1 \leq T < 5$ 的格子数不超过 $1/5$ 格子总数，查阅皮尔逊卡方

卡方检验的时候多重检验法有卡方分割法和 bonferroni 方法，卡方分割法在 SPSS 中没法自动实现，bonferroni 方法仅限于比较列之间的比例，因此要比较不同地区人群 ABO 血型分类构成比的时候，在交叉表格中行选择血型，列选择地区，并在上述操作中，点击单元格，勾选 z 检验比较列的比例，并勾选调整 p 值 (Bonferroni 方法)，在结果中会用 abc 字母下标加以区分。

3. 拟合优度检验

3.1 构成比的拟合优度检验

试比较 1.3 例题中欧洲与北美洲的 ABO 血型构成比是否相同

对 1.3 的 SPSS 数据进行个案到变量的重组 (个案重组为变量，纵向数据转变为横向数

据)，标识变量为血型，索引变量为地区，结果图如下

血型	亚洲	欧洲	北美洲
A型血	321	258	408
B型血	369	43	106
AB型血	95	22	37
O型血	295	194	444

简易操作步骤如下：

加权个案，加权变量为欧洲

分析 → 非参数检验 → 旧对话框 → 卡方 → 检验变量列表为血型，期望值将北美洲的人数数值从上到下逐个输入 → 确定

$H_0: \pi_1 = \pi_2$ ，不能认为欧洲与北美洲的 ABO 血型构成比有显著差异

$H_1: \pi_1 \neq \pi_2$ ，可以认为欧洲与北美洲的 ABO 血型构成比有显著差异

$\alpha=0.05$

输出结果：

血型			
	实测个案数	期望个案数	残差
A型	258	212.0	46.0
B型	43	55.1	-12.1
AB型	22	19.2	2.8
O型	194	230.7	-36.7
总计	517		

检验统计	
	血型
卡方	18.871 ^a
自由度	3
渐近显著性	.000
a. 0 个单元格 (.0%) 的期望频率低于 5。期望的最低单元格频率为 19.2。	

$P=0.000<0.05$ ，拒绝 H_0 ，接受 H_1 ，因此可以认为欧洲与北美洲的 ABO 血型构成比有显著差异。

3.2 二项分布的拟合优度检验

将某批菌种在特定的培养基中进行保存，共 60 盒，每盒 10 支，在液氮中保存相当一段时间后，因实验需要将菌种复苏，每盒菌种的活性情况如下表，试检验菌种的活性存活是否服从二项分布？

失活数	观测到次数
0	8
1	15
2	20
3	10
4	5
5	2

操作步骤如下：

1. 计算菌种的失活概率。
观测值 $p = (0 \times 8 + 1 \times 15 + 2 \times 20 + 3 \times 10 + 4 \times 5 + 5 \times 2) / (60 \times 10) = 0.1917$
2. 在 SPSS 中计算二项式分布的理论概率
转换→计算变量→PDF 与非中心 PDF→PDF.BINOM(菌种的失活数,10,0.1917)
3. 以“观测次数”加权个案
4. 卡方检验中以菌种失活数目为检验变量，在期望值中输入二项分布的理论频率，分析实际观测数值与理论预期值之间的差异
5. 思考：做拟合优度检验时失活数是否需要覆盖 0~10 所有可能性？

H0: 可以认为菌种活性存活服从二项分布

H1: 不可以认为菌种活性存活服从二项分布

$\alpha=0.05$

输出结果：

失活菌种			
	实测个案数	期望个案数	残差
0	8	7.3	.7
1	15	17.0	-2.0
2	20	18.2	1.8
3	10	11.5	-1.5
4	5	4.8	.2
5	2	1.2	.8
总计	60		

检验统计	
失活菌种	
卡方	1.199 ^a
自由度	5
渐近显著性	.945
a. 2 个单元格 (33.3%) 的期望 频率低于 5。期 望的最低单元格 频率为 1.2。	

$P=0.945>0.05$ ，拒绝 H1，接受 H0，所以可以认为菌种活性存活服从二项分布。

思考：失活数需要覆盖 0~10 所有可能性。但实际操作中，可以进行合并，例如将 ≥ 5 的合并进行拟合优度检验。

3.3 思考题：题 2.1 可否使用拟合优度检验？如可以，p 值是否相同？为什么？

可以。

方法：先手动计算各个部分的预期值，然后执行拟合优度检验。

drug	effect	frequency	predicted
口服	有效	58	61.95
口服	无效	40	36.05
注射	有效	64	60.05
注射	无效	31	34.95

H0：不能认为两种给药方式效果有显著差异

H1：可以认为两种给药方式效果有显著差异

$\alpha=0.05$

输出结果：

frequency			
	实测个案数	期望个案数	残差
注射 无效	31	35.0	-4.0
口服 无效	40	36.1	4.0
口服 有效	58	62.0	-4.0
注射 有效	64	60.1	3.9
总计	193		

检验统计	
frequency	
卡方	1.391 ^a
自由度	3
渐近显著性	.708
a. 0 个单元格 (0.0%) 的期望 频率低于 5。期 望的最低单元格 频率为 35.0。	

$P=0.708>0.05$ ，拒绝 H1，接受 H0，不能认为两种给药方式有显著差异。此法与卡方检验方法相比，自由度不同：拟合优度检验自由度为 3，卡方检验自由度为 1；同时两法得到的 p 值也不同，卡方检验的 $p=0.238$ 。