

【例题】相关系数检验力的计算，我们首先应该计算效应量的大小(d)，对于检验两个变量之间的相关系数是否等于 0，定义：**★ ρ 是相关系数**

$$d = \rho_1 - \rho_0 = 0.3$$

$\rho_1 = 0.3$ 对应于备选假设 H_1 ， $\rho_0 = 0$ 对应的是虚无假设 H_0 。

$$\text{定义 } \delta = d \sqrt{n-1} = \rho_1 \sqrt{n-1} \rightarrow \delta = r \cdot \sqrt{n-1}$$

对于样本量 $n=50$ ，可以计算得到：

$$\delta = d \sqrt{n-1} = \rho_1 \sqrt{n-1} = 0.3 \times \sqrt{50-1} = 2.1$$

对于 $\delta=2.1$ 和 $\alpha=0.05$ (双侧)，**查检验力表(附表 4)** 可以得到检验力

测定系数 R^2 ：(非测定系数 $1-R^2$)

变异比例才占 50%。以上举例表明 r 仅说明 X 变量和 Y 变量的关系程度， r 才能说明 X 变量对 Y 变量所能解释的比例。例如，吸烟与否和癌症患者的相关 $r=$

参数估计：

总体均值(或总体均值差异) = 样本均值(或均值差异) $\pm t$ (估计标准误)

对于单样本的 t 统计量，我们将使用一个样本均值 $\bar{X}$ ，来估计总体均值 μ 。单样本中，总体均值的估计公式如下：

$$\mu = \bar{X} \pm tS_{\bar{X}} \quad (11-3)$$

对于独立测量的 t 统计量，我们将使用一个样本均值差异 $\bar{X}_1 - \bar{X}_2$ ，来估计总体均值差异 $\mu_1 - \mu_2$ 。独立测量中，总体均值差异的估计公式如下：

$$\mu_1 - \mu_2 = (\bar{X}_1 - \bar{X}_2) \pm tS_{\bar{X}_1 - \bar{X}_2} \quad (11-4)$$

最后，重复测量的 t 统计量可以使用样本均值差异 $\bar{X}_D$ 来估计总体均值差异 μ_D 。重复测量中，总体均值差异公式如下：

$$\mu_D = \bar{X}_D \pm tS_{\bar{X}_D} \quad (11-5)$$

单因素：

变异来源	SS	df	MS	F	P
处理间	70	2	35	9.14	<0.05
处理内	46	12	3.83		
合计	116	14			

键值 A	键值 B	键值 C
0	5	5
1	4	6
0	2	5
1	6	0
2	8	4
2	8	4
$T_i=5$	$T_j=25$	$T_k=30$
$n_i=5$	$n_j=5$	$n_k=5$

$$SS_{\text{总}} = \sum X^2 - \frac{G^2}{N} = 356 - \frac{60^2}{15} = 116$$
$$SS_{\text{处理间}} = \sum \frac{T_i^2}{n_i} - \frac{G^2}{N} = \frac{5^2}{5} + \frac{25^2}{5} + \frac{30^2}{5} - \frac{60^2}{15} = 70$$
$$SS_{\text{处理内}} = SS_{\text{总}} - SS_{\text{处理间}} = 116 - 70 = 46$$
$$df_{\text{总}} = N - 1 = 15 - 1 = 14$$
$$df_{\text{处理间}} = k - 1 = 3 - 1 = 2$$
$$df_{\text{处理内}} = df_{\text{总}} - df_{\text{处理间}} = N - k = 14 - 2 = 12$$

重复测量：

误差来源	SS	df	MS	F	P
处理间	28.35	3 k-1	9.52	41.39	<0.05
处理内	18.4	16			
被试间	15.7	4 n-1			
误差	2.7	12	0.23		
合计	46.95	19 nk-1			

被试	实验处理				P	N=20
	12小时	24小时	36小时	48小时		
1	5	4	3	2	14	$G=99$ $\sum X^2=537$
2	6	5	5	3	19	
3	7	6	4	3	20	
4	8	7	6	4	25	
5	6	6	5	4	21	
$T_i=32$	$T_j=28$	$T_k=23$	$T_l=16$			
$n_i=5$	$n_j=5$	$n_k=5$	$n_l=5$			

式14-6, T_i 就变成了现在的 P_i , n 就变成了现在的 k , 而 N 和 G 保持不变。这样可以得到：

$$SS_{\text{处理间}} = \sum \frac{P_i^2}{k} - \frac{G^2}{N}$$

两因素：

变异来源	SS	df	MS	F	P
处理间	413.60	5 3*2-1			
群体性质	31.69	1 2-1	31.69	6.47	$P<0.05$
群体人数	325.29	2 2-1	162.65	33.19	$P<0.001$
群体性质×群体人数交互作用	56.62	2	28.31	5.78	$P<0.05$
处理内	205.88	42 N-2			
合计	619.48	47 N-1			

群体性质		群体人数		N=48
真实群体	虚拟群体			
0	$T=143$ $n=8$	$T=137$ $n=8$	$T_{ij1}=280$ $n_{ij1}=16$	$G=685$
	$T=109$ $n=8$	$T=118$ $n=8$	$T_{ij2}=227$ $n_{ij2}=16$	
1	$T=71$ $n=8$	$T=107$ $n=8$	$T_{ij3}=178$ $n_{ij3}=16$	$\sum X^2=10395$
	$T_{i11}=323$ $n_{i11}=24$	$T_{i12}=362$ $n_{i12}=24$		

$$SS_{\text{总}} = \sum X^2 - \frac{G^2}{N} = SS_{\text{处理间}} = \sum \frac{T_{ij}^2}{n_{ij}} - \frac{G^2}{N}$$

$$SS_{\text{群体人数}} = \sum \frac{T_{i1}^2}{n_{i1}} - \frac{G^2}{N}$$
$$SS_{\text{群体性质}} = \sum \frac{T_{i1}^2}{n_{i1}} - \frac{G^2}{N}$$

测量效应：

公式：

$$\eta^2/r^2 = \frac{df_{\text{处理间}} \times F}{df_{\text{处理间}} \times F + df_{\text{处理内}}}$$

$$\eta^2/r^2 = \frac{SS_{\text{处理间}}}{SS_{\text{总}}} = \frac{SS_{\text{处理间}}}{SS_{\text{处理间}} + SS_{\text{处理内}}}$$

的计算方法为：

Cohen's f $f = \sqrt{\frac{\eta^2}{1-\eta^2}}$

事后检验：

为两个处理组之间的差异不显著。HSD 临界值的计算公式如下：

如果 $|x_i - x_j| > HSD = q \sqrt{\frac{MS_{\text{误差}}}{n}}$, 则差异显著。

从总变异中去除, 然后来计算 η^2 (响应的)

重复测量单因素方差分析中计算测量效应的公式如下：

$$\eta^2 = \frac{SS_{\text{处理间}}}{SS_{\text{总}} - SS_{\text{被试间}}} = \frac{SS_{\text{处理间}}}{SS_{\text{误差}} + SS_{\text{处理内}}}$$

Cohen's $d = \frac{x_i - x_j}{\sqrt{MS_{\text{误差}}}}$

$$HSD = q \sqrt{\frac{MS_{\text{误差}}}{n}}$$

数量 $k=4$, 误差自由度 $df_{\text{误差}}=c$
在显著性水平为 0.05 的假设下
临界值。

$$\eta^2_{\text{群体性质}} = \frac{SS_{\text{群体性质}}}{SS_{\text{总}}} = \frac{31.69}{619.48} = 0.051$$
$$\eta^2_{\text{群体人数}} = \frac{SS_{\text{群体人数}}}{SS_{\text{总}}} = \frac{325.29}{619.48} = 0.525$$
$$\eta^2_{\text{群体性质} \times \text{群体人数}} = \frac{SS_{\text{群体性质} \times \text{群体人数}}}{SS_{\text{总}}} = \frac{56.62}{619.48} = 0.091$$
$$f_{\text{群体性质}} = \sqrt{\frac{\eta^2_{\text{群体性质}}}{1-\eta^2_{\text{群体性质}}}} = \sqrt{\frac{0.051}{1-0.051}} = 0.23$$
$$f_{\text{群体人数}} = \sqrt{\frac{\eta^2_{\text{群体人数}}}{1-\eta^2_{\text{群体人数}}}} = \sqrt{\frac{0.525}{1-0.525}} = 1.05$$
$$f_{\text{群体性质} \times \text{群体人数}} = \sqrt{\frac{\eta^2_{\text{群体性质} \times \text{群体人数}}}{1-\eta^2_{\text{群体性质} \times \text{群体人数}}}} = \sqrt{\frac{0.091}{1-0.091}} = 0.32$$

而, 我们可以运用相同的方法, 计算主效

使用 Tukey 的 HSD 检验法进行说明。

“群体人数”的事后检验

对于这个因素, 处理组的数量为 3, 处理内自由度为 42。我们首先在附表 8 得到在显著性水平为 0.05 的假设下, q 的临界值为 3.44。然后, 根据前面的计算, 我们已经得到处理内均方为 $MS_{\text{误差}}=4.90$, 每个因素水平上的被试人数为 16。这样, 就能计算出 HSD 临界值为：

$$HSD = q \sqrt{\frac{MS_{\text{误差}}}{n}} = 3.44 \times \sqrt{\frac{4.90}{16}} = 1.90$$

接下来的步骤就是计算任意两个样本的均值差异, 并与 HSD 临界值做比较。
 $|M_{0.0.0.0.0.} - M_{0.0.0.0.0.}| = |17.50 - 14.19| = 3.31 > 1.90$
 $|M_{0.0.0.0.0.} - M_{0.0.0.0.0.}| = |17.50 - 11.13| = 6.37 > 1.90$
 $|M_{0.0.0.0.0.} - M_{0.0.0.0.0.}| = |14.19 - 11.13| = 3.06 > 1.90$

因此我们可以得出结论, 在 0.05 显著性水平上, 群体人数任意两个水平之间的差异都显著, 这也证明了社会催化效应随着群体人数的增加, 而显著增加。

交互作用的事后检验

两个因素交互之后一共得到 6 种处理水平的结合, 如果对所有可能的两组合进行事后检验, 一共有 $C_6^2=15$ 组事后检验。因此, 出于平均时间的考虑, 我们往往根据研究目的的需要, 选择一些感兴趣的进行事后比较。

首先, 我们还计算出交互作用事后检验的 HSD 临界值。(注意这时查表 q 临界值用到的信息变为 $k=6$, $df_{\text{误差}}=42$, 并且每个处理水平结合的被试人数为 8;)

$$HSD = q \sqrt{\frac{MS_{\text{误差}}}{n}} = 4.23 \times \sqrt{\frac{4.90}{8}} = 3.31$$

然后, 我们选择几对有意义的处理水平的结合, 来进行多重比较, 见表 15-6 和表 15-7。

表 15-6 例 15.2 的事后检验表 I			
	0	1	5
真实群体-虚拟群体 I	0.75<3.31	1.13<3.31	4.50>3.31

皮尔逊相关：两连续变量的线性关系， $n \geq 30$

二、协方差 $\rightarrow$ 大小跟数据单位数²⁰同

前面所讲述的平均值描述的是一组数据的集中趋势，标准差、方差描述的是
一组数据的离散程度，而协方差描述的是两个随机变量之间的关系，方差是协方
差的一种特殊情况，即两个变量是相同的情况。例如，学生的自我效能感和学业
成绩之间的关系。协方差的计算公式如下：

$$\text{cov} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{n - 1} \quad (16-1)$$

公式中： x, y 同高同低

$\bar{X}$ 为变量 X 的平均数； $\bar{Y}$ 为变量 Y 的平均数； n 为数据的对数。

如果协方差为正，则说明两变量之间呈现正相关关系，也就是说学生自我效能感
越高，学生学业成绩越好；如果协方差为负，说明两变量之间呈现负相关关系，也就
是说学生自我效能感越高，学生学业成绩越差；如果协方差为零，说明两变量之间呈
现零相关关系，也就是说学生自我效能感和学生的学业成绩没有关系。

三、皮尔逊相关的计算公式 x, y 同高同低

《连续变量的线性相关》《将体育成绩标准化》

皮尔逊相关系数用符号 r 表示，是 X 和 Y 共变程度与 X 和 Y 变异程度的比
值，也是 X 和 Y 的协方差与 X 和 Y 各自的标准差的比值。其计算公式如下：

$$r = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}} = \frac{\text{cov}}{s_x \cdot s_y} \quad (16-2)$$

T检验检验：

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} \quad df = n - 2$$

$$r = \frac{\sum XY - (\sum X)(\sum Y)/n}{\sqrt{\sum X^2 - (\sum X)^2/n} \times \sqrt{\sum Y^2 - (\sum Y)^2/n}}$$

斯皮尔曼等级相关：两等级变量，不要求 $n \geq 30$

公式不复杂。

曼等级相关的计算公式 排序的排列。

★ 最大等级为1

曼等级相关系数表示为 r_s ，当使用皮尔逊相关对等级数据进行计算
为斯皮尔曼相关，因此，有人认为斯皮尔曼等级相关是皮尔逊相关的
其还可以将等级直接用于更简单的公式，其计算公式为：排序数的平方

T检验检验：查表

$$r_s = 1 - \frac{6 \sum (R_{xi} - R_{yi})^2}{n(n^2 - 1)} = 1 - \frac{6 \sum D^2}{n(n^2 - 1)} \quad (16-10)$$

$$C = \frac{m(m^2 - 1)}{12}$$

m ：变量中某个相同等级出现的次数

公式中 C 为校正数， m 为相同等级的数目。

如果出现相同等级，斯皮尔曼等级相关的计算公式为：

$$r_{RC} = \frac{\sum x^2 + \sum y^2 - \sum D^2}{2 \times \sqrt{\sum x^2 \times \sum y^2}}$$

$$\sum x^2 = \frac{N^3 - N}{12} - \sum C_x$$

$$\sum y^2 = \frac{N^3 - N}{12} - \sum C_y$$

肯德尔和谐系数: r_w 越大越一致

肯德尔和谐系数的计算公式:

$$r_w = \frac{SS_R}{\frac{1}{12}K^2(n^3 - n)}$$

其中,

$$SS_R = \sum R_i^2 - \frac{(\sum R_i)^2}{n} = \sum (R_i - \bar{R})^2$$

公式中: r_w 表示肯德尔和谐系数;

K 表示等级评定者的数目, 即变量数;

n 为被等级评定的对象的数目;

R 为被评定对象获得的 K 个等级之和;

SS_R 为 R 的离差平方和。

类似斯皮尔曼等级相关公式, 有相同等级时, 需要用肯德尔和谐系数的修正

公式计算 r_w 系数。

肯德尔和谐系数的修正公式如下:

$$r_w = \frac{SS_R}{\frac{1}{12}K^2(n^3 - n) - K \sum C} \quad (16-18)$$

其中, $\sum C = \sum \frac{t^3 - t}{12}$, t 为相同等级的数目。

【例 16.6】

四个评委给五个选手

假设检验: 查表12

点二列相关: 连续——真二分

$$r_{pb} = \frac{\bar{X}_1 - \bar{X}_2}{S_t} \cdot \sqrt{\frac{n_1 n_2}{n(n-1)}}$$

$\bar{X}_1$ 是与二分名义变量的一个值对应的连续变量的平均数;

$\bar{X}_2$ 是与二分名义变量的另一个值对应的连续变量的平均数;

n_1 和 n_2 分别是二分名义变量的两个值各自对应的数据个数, $n_1 + n_2 = n$;

S_t 是连续变量对应的样本标准差。

假设检验:

$$t = \frac{r_{pb}}{\sqrt{\frac{1 - r_{pb}^2}{N - 2}}} \quad df = N - 2$$

二列相关: 连续——人为二分

公式中, Y 为标准正态分布曲线中与 p 值对应的纵线高度; $r_b = \frac{\bar{X}_p - \bar{X}_q}{S_t} \times \frac{pq}{Y}$ 其余符号与点二列相关计算公式中含义相同。

$\bar{X}_p$ 是与二分名义变量的一个值对应的连续变量的平均数;

$\bar{X}_q$ 是与二分名义变量的另一个值对应的连续变量的平均数;

p 和 q 分别是二分名义变量的两个值各自所占的比例, $p + q = 1$;

S_t 是连续变量的标准差。

$$r_b = \frac{\bar{X}_p - \bar{X}_q}{S_t} \times \frac{p}{Y}$$

用的是 Z 检验,

$$Z = \frac{r_b}{\frac{1}{y} \times \sqrt{\frac{pq}{n}}}$$

Φ 相关: 二分——二分

Φ 相关系数计算公式为:

$$r_\phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} \quad (16-25)$$

● 列联性检验 ●

● 列联性检验: 变量的不同取值是否属于不同的类别。

● 期望频数: $E = \frac{(a+b)(c+d)}{n}$

● 零假设: $H_0: a+b=c+d$

● $K^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$

● 列联性检验: 用列联性检验判断两个分类变量是否独立。

反证法原理: 在假设 H_0 下, 如果推出一个矛盾, 就证明了 H_0 不成立。

独立性检验原理: 在假设 H_0 下, 如果出现一个与 H_0 相矛盾的小概率事件, 就推断 H_0 不成立, 且该推断犯错误的概率不超过这个概率。

$P(K^2 \geq k_0)$	0.50	0.40	0.25	0.10	0.05	0.025	0.010	0.005	0.001
k_0	0.455	0.708	1.323	2.072	2.706	3.841	5.024	6.635	10.828

若计算出的 K^2 值 $< k_0$, 则推断“X与Y没有关系”错误的概率不超过 $P(K^2 \geq k_0)$ 。

Φ 相关的假设检验

Φ 系数计算完之后, 同样需要检验其显著性。Φ 系数显著性检验的虚无假设为“总体中不存在相关”; 备择假设为“总体中存在相关”。该检验是一个卡方检验, 因为 Nr_ϕ^2 服从自由度为 1 的卡方分布。

$$K^2 = r^2 \cdot n$$

回归

回归系数: $a = \bar{y} - b\bar{x}$

$$b = \frac{SP}{SS_x} = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} = r \times \frac{S_y}{S_x} \quad (\text{标准化的}\beta\text{同理})$$

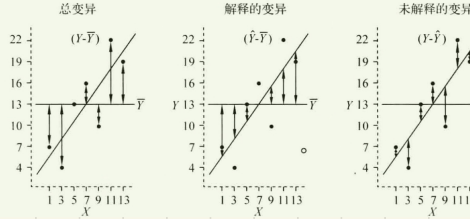
变异分解:

各点的Y到其平均值Y的距离都可分为两部分:

$$(Y - \bar{Y}) = (Y - \hat{Y}) + (\hat{Y} - \bar{Y})$$

(17-7)

如果各点都很接近回归线, 则各点的Y值到其平均值的离差中大部分是由回归部分的差异造成的, 这种情况说明回归误差小, 回归的效率较高。



$$\sum (Y - \bar{Y})^2 = \sum (\hat{Y} - \bar{Y})^2 + \sum (Y - \hat{Y})^2$$

$SS_T = SS_R + SS_E$
 $SS_Y = SS_{回归} + SS_{残差}$

$$R^2 = \frac{SSR}{SST} = b_1^2 \frac{SS_X}{SS_Y} = b_1^2 \frac{S_X^2}{S_Y^2} = r^2 \quad (\text{一元回归中})$$

方程有线性 ← (等价) → 回归系数 b 显著性

标准误:

对于回归方程, 其估计标准误的计算公式为:

$$S_{YX} = \sqrt{\frac{\sum (Y - \hat{Y})^2}{n - 2}} = \sqrt{\frac{SS_E}{df}} \quad (17-8)$$

其中, S_{YX} 是估计标准误。

计算上述公式较为简单的方法是:

$$S_{YX} = S_y \sqrt{(1 - r^2) \frac{n-1}{n-2}} \quad (17-9)$$

注意: 如果 $r=1.0$ 或者 $r=-1.0$, $S_{YX}=0$, 也就是说预测值和回归直线没有任何偏差, 而当 $r=0$ 时, 误差最大, S_{YX} 几乎等于 S_y 。即 $S_{YX} = S_y \sqrt{\frac{n-1}{n-2}}$, 如果样本量是足够大, 则 S_{YX} 可以近似等于 S_y 。

在回归直线上, 当与所有自变量 X 相对应的各组因变量 Y 的残差值都服从正态分布, 并且残差方差齐性时, 由 X 估计 Y 的回归系数的标准误为:

$$SE_b = \frac{SE_{YX}}{S_x \sqrt{\sum (X - \bar{X})^2}} \quad (17-13)$$

SE_{YX} 表示对 Y 进行估计的标准误。

当样本容量较大, 回归系数估计的标准误可以表示为:

$$SE_b = \frac{S_y}{S_x} \sqrt{\frac{1 - r^2}{n - 2}} \quad (17-14)$$

式中 S_y 是变量 Y 的样本标准差, S_x 是变量 X 的样本标准差, r 是变量 X 与 Y 的相关系数。

表 17-1 回归方程方差分析表

变异来源	平方和	自由度	方差	F 值	概率
回归	SS_R	$df_R = 1$	MS_R	$F = \frac{MS_R}{MS_E}$	P
残差	SS_E	$df_E = n - 2$	MS_E		
总变异	SS_T	$df_T = n - 1$			

上述平方和可以用原始数据直接计算:

$$SS_T = \sum (Y - \bar{Y})^2 = \sum Y^2 - \frac{(\sum Y)^2}{n}$$

$$SS_R = \sum (\hat{Y} - \bar{Y})^2 = b^2 \cdot \left[\sum X^2 - \frac{(\sum X)^2}{n} \right]$$

其中 $SS_E = SS_T - SS_R$

具体检验步骤如下:

(1) 提出假设

$$H_0: \beta = 0, H_1: \beta \neq 0$$

(2) 计算检验统计量的值

$$t = \frac{b - \beta}{SE_b}$$

$$\text{式中 } SE_b = \sqrt{\frac{S_{YX}^2}{\sum (x - \bar{x})^2}}, S_{YX} = \sqrt{\frac{\sum (Y - \hat{Y})^2}{n - 2}}$$

(3) 对于给定的显著性水平 α , 查自由度为 $n-2$ 的双侧 t 值表,

$$F = t^2$$

计数数据的检验

二项检验:

$n < 50$: 查表

$n \geq 50 / np \geq 10$: Z检验

对二项分布的随机事件发生次数的检验, 其对应的统计量为:

$$Z = \frac{m - np}{\sqrt{npq}}$$

其中, m 为二项分布所关心的随机事件 A 实际发生的实际次数,

- 如果两个样本中的比率 $\hat{p}_1$ 和 $\hat{p}_2$ 分别表示两个样本中同一事件发生的比率, 且 $n_1\hat{p}_1 \geq 10$, $n_2\hat{p}_2 \geq 10$ (n_1, n_2 分别代表两样本的样本量) 时, 如果假设是 $\hat{p}_1$ 和 $\hat{p}_2$ 的差等于一个不为 0 的值, 其标准误的计算公式为:

$$SE_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{(n_1\hat{p}_1 + n_2\hat{p}_2)(n_1\hat{q}_1 + n_2\hat{q}_2)}{n_1n_2(n_1 + n_2)}}$$

- 统计量的计算公式为:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{SE_{\hat{p}_1 - \hat{p}_2}}$$

- 然后确定显著性, 查标准正态分布表确定差异是否显著, 对结果进行解释。

卡方检验:

拟合度:

配合度的卡方检验是一种主要用于检验单一变量(一个因素多项分类)的实际观察次数与理论次数是否存在差异或是否符合某一比例的推断统计方法。

$H_0: f_o = f_e$ 或者 $f_o - f_e = 0$; $H_1: f_o - f_e \neq 0$

① 一般情况下:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

② 次数较小(小于 5)时的修正

$$\chi^2 = \sum \frac{(f_o - f_e - 0.5)^2}{f_e}$$

χ^2 检验自由度的确定: 一般为 $df = k - 1$, k 为分类数目。

独立性:

数据就是四格表, 如果一个因素有 R 类, 另外一个因素有 C 类, 则为 $R \times C$ 表。

独立性检验是主要研究两个分类变量之间是否存在关联的一种 χ^2 检验。

$$f_e = \frac{f_{.j} f_{.i}}{N}, df = (R-1)(C-1), \chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

对于四格表的独立性检验, 相当于独立样本比率差异的显著性检验。当每个单元格的期望次数大于等于 5 时, 也可以使用下面的简便公式计算 χ^2 值。

$$\chi^2 = \frac{N(AD - BC)^2}{(A+B)(C+D)(A+C)(B+D)} \quad (18-7)$$

χ 相关统计量

φ 相关计算公式为:

$$\phi = \sqrt{\frac{\chi^2}{n}}, \chi^2 = n\phi^2$$

χ^2 是四格表独立性检验所计算的 χ^2 值。

另外对于四格表(独立样本)相关程度的描述, 除常用 φ 相关外, 有时还用到其他表示方法。例如尤尔的其他系数:

关联系数

$$Q = \frac{ad - bc}{a+b+c+d}$$

相关系数(有实用价值)

$$\gamma = \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}}$$

(18-10)

这些表示 χ^2 表计数资料二因素之间相关程度的尺度不同, 数值也可能不同, 但都能反映两因素之间的相关。Q 与 γ 之间还有一定的数量关系:

$$Q = 2\gamma(1 + \gamma^2) \quad (18-11)$$

$R \times C$ 的列联相关:

$$\text{列联系数 } C = \sqrt{\frac{\chi^2}{n + \chi^2}}$$

$$T = \sqrt{\frac{\chi^2}{df + (c-1)N}}$$

$$\text{Cramer's } V = \sqrt{\frac{\chi^2}{n(df^*)}}, (df^*) = \min(R-1, C-1)$$

$$W = \sqrt{\frac{\chi^2}{n}}$$

H_0 : 观测数据服从正态分布。
 H_1 : 观测数据不服从正态分布。
 计算 $p=80, \alpha=6, n=35$
 用频数分布表, 并计算非参数函数 $F_n(x)$ 和理论分布函数 $F(x)$, 从而
 计算出 D 值:

观测频数 $f(x_i)$	次数 i	累计频数 $F_n(x)$	理论函数 $F(x)$	绝对差 $ F_n(x) - F(x) $	理论分布 $F(x)$	$D = \max_i F_n(x) - F(x) $
68	2	2	0.0531	-2.00	0.0228	0.1201
72	2	4	0.1143	-1.33	0.0534	0.0209
75	2	6	0.1714	-0.83	0.1033	0.0333
76	2	8	0.2286	-0.67	0.2514	0.0228
77	6	14	0.4000	-0.50	0.3085	0.0912
78	3	17	0.4857	-0.23	0.3707	0.1100
80	6	23	0.6571	0	0.5000	0.1371
81	3	26	0.7429	0.17	0.5675	0.1356
84	2	28	0.8000	0.87	0.7486	0.0514
86	2	30	0.8571	1.00	0.8433	0.0358
87	2	32	0.9143	1.17	0.8789	0.0352
89	3	35	1.0000	2.00	0.9772	0.0228

 $D = \max_i |F_n(x) - F(x)| = 0.1714, D(35, 0.05) = 0.224$, 因此 $D < D^*$
 (0.05, 0.05), 接受零假设, 即该观测数据服从正态分布。

非参数

单样本

k-S 检验

(比较与指定分布的匹配度)
 < 常为正态分布 >

两独立

U 检验

(比较两个样本分布是否相同)

中位数检验

(比较两个样本对应的总体的中位数是否相同)

多独立

Kruskal-Wallis H 检验

(比较多个样本分布是否相同, 是 U 检验在多样本的推广)

中位数检验 (同上)

两相关

卡方检验

$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$Z = \frac{\sqrt{n}(\tau - 0.5) - \frac{N}{2}}{\sqrt{\frac{N}{2}}}$$
 式中: τ 为 $\frac{\min(n_i, n_j)}{\min(n_i, n_j)}$ 的数值, N 为 n_i 与 n_j 之和。

符号检验

小样本时:
 $r < r_0$ 时拒绝 H_0

(比较两个配对样本的分布是否相同)

Wilcoxon 符号秩检验

(同上, T 用于原始数据, 结果更准确)

弗里德曼检验

(比较多个配对样本的分布是否相同)

多相关

Wilcoxon 符号秩检验的计算步骤:
 1. 小样本 ($n \leq 25$) 时的检验方法
 首先, 按照符号检验的方法, 分别用第二组样本的各个观测值减去第一组对应样本的观测值。差值为正则记为符号, 为负则记为符号, 并同时保存差值数据; 然后, 将差值绝对值按升序排序, 如果差值为零时, 零不参加等级排列; 最后, 分别计算正符号秩和 W^+ 和负符号秩和 W^- , 取两者中之较小的记为 W 。
 2. 大样本 ($n > 25$) 时的检验方法 $W \sim N(\frac{n(n+1)}{4}, \frac{n(n+1)}{24})$
 一般认为 W 的分布接近正态分布, 其平均数为:

$$\mu_W = \frac{n(n+1)}{4}$$

$$\sigma_W = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$
 因而可以进行 Z 检验:

$$Z = \frac{W - \mu_W}{\sigma_W}$$

第一步, 将每一被试 (或匹配对象) 的 k 个数据从小到大排列, 若遇到数据相等则取这几个数据所在排列的平均。
 第二步, 计算每组成备给下 n 个被试的秩数, 记为 $r_i, i=1, 2, \dots, k$ 。
 第三步, 代入公式 (19-42), 计算 T 统计量为秩和的统计量:

$$T = \frac{12}{n(n+1)} \sum_{i=1}^k R_i^2 - 3n(k+1) \quad (19-42)$$
 第四步, 作出决策, 这里可以分两类情况:
 (1) 样本容量较小或数据较少情况
 附表 19 给出了样本容量较小, 数据较少下各显著性检验的临界值, 在这种情况下, 可在计算出 T 后, 根据设定的显著性水平 α , 将样本容量 n 和显著性水平 α 在表中查找临界值, 如果 T 大于表中临界值, 则拒绝原假设, 否则无法拒绝原假设。
 (2) 样本容量较大或数据较多情况
 当数据较多, 样本量较大时, 难以查找到临界值, 在这种情况下, T 统计量近似服从