

## 定性数据统计分析作业 (6)

钟瑜 222018314210044

2020 年 12 月 11 日

1. 三家供应商提供的零件合格和不合格的情况如下:

供应商	零件质量		
	良好	小缺陷	大缺陷
A	90	3	7
B	170	18	7
C	135	6	9

取  $\alpha = 0.05$ , 检验供应商与零件质量的独立性。你的分析结果能告诉采购部门什么?

解.

```
1 > x<-matrix(c(90,170,135,3,18,6,7,7,9),nrow=3)
2 > chisq.test(x,correct = F)
3
4 Pearson's Chi-squared test
5
6 data:  x
7 X-squared = 7.7117, df = 4, p-value = 0.1027
```

p 值大于  $\alpha=0.05$ , 故接受原假设, 认为供应商与零件质量独立, 即选哪个供应商, 其质量都差不多

2. 向 100 个女性和 100 个男性作调查, 了解他们关于给谁买节日礼物最难的看法。调查结果如下:

性别	给谁买礼物最难					
	配偶	父母	子女	兄弟姐妹	姻亲	其他亲属
女性	25	31	19	3	10	12
男性	37	28	7	8	4	16

女性和男性关于给谁买节日礼物最难的看法上有没有显著的差异?

解.

```
1 > x2<-matrix(c(25,37,31,28,19,7,3,8,10,4,12,16),nrow=6)
2 > chisq.test(x2,correct = F)
```

```

3
4 Pearson's Chi-squared test
5
6 data:  x2
7 X-squared = 32.927, df = 5, p-value = 3.892e-06

```

p 值于小  $\alpha=0.01$ , 故不接受原假设, 认为女性和男性关于给谁买节日礼物最难的看法上有显著差异.

3. 调查人们对某项措施的满意程度, 可以问他:“你对这项措施满意吗?”, 也可以问他:“你对这项措施不满意吗?”为了解这两种提问方式对被调查者回答问题有没有影响, 向 243 人问:“你满意吗?”, 另外向 240 人问:“你不满意吗?”. 调查结果如下:

		提出的问题	
		你满意吗	你不满意吗
回答	非常满意	139	128
	比较满意	82	69
	比较不满意	12	20
	不满意	10	23

问:这两种提问方式对被调查者回答问题有没有影响?

解.

```

1 > x3<-matrix(c(139,82,12,10,128,69,20,23),nrow=2)
2 > chisq.test(x3,correct = F)
3
4 Pearson's Chi-squared test
5
6 data:  x3
7 X-squared = 5.705, df = 3, p-value = 0.1269

```

p 值大于  $\alpha=0.01$ , 故接受原假设, 认为这两种提问方式被调查者回答问题没有影响.

5. 278 例尸体解剖资料整理如下

年龄(岁)	冠状动脉硬化等级(由低到高)				合计
	-	+	++	+++	
20 至 30	70	22	4	2	98
30 至 40	27	24	9	3	63
40 至 50	16	23	13	7	59
50 以上	9	20	15	14	58
合计	122	89	41	26	278

年龄越大的人,冠状动脉硬化的程度是否有越重的趋势?(取水平  $\alpha = 0.05$ )

解.

```

1 > congruence_test=function(x,alternative="twoside")
2 +   #适用于列联表的相合性检验问题
3 +   #x为列联表矩阵; alternative对应于备择假设
4 + {
5 +   n=sum(x)
6 +   G=0;H=0
7 +   r=nrow(x)
8 +   c=ncol(x)
9 +   r1=r-1;c1=c-1
10 +   for (i in 1:r1){
11 +     for (j in 1:c1){
12 +       G=G+x[i,j]*sum(x[(i+1):r,(j+1):c])
13 +     }
14 +   }
15 +   for (i in 1:r1){
16 +     for (j in 2:c){
17 +       H=H+x[i,j]*sum(x[(i+1):r,1:(j-1)])
18 +     }
19 +   }
20 +   z=G-H
21 +   TA=sum(rowSums(x)*(rowSums(x)-1)/2)
22 +   TB=sum(colSums(x)*(colSums(x)-1)/2)
23 +   #TAB=G+H+TA+TB-n*(n-1)/2
24 +   Cn2=n*(n-1)/2
25 +   #计算各系数的值
26 +   Kendall_TAO=z/sqrt((Cn2-TA)*(Cn2-TB))
27 +   Gamma=(G-H)/(G+H)
28 +   d_BA=(G-H)/(Cn2-TA)
29 +   d_AB=(G-H)/(Cn2-TB)

```

```

30 +
31 +     #近似公式,表示sigma的平方
32 +     sigma_2=(n^3-sum(rowSums(x)^3))*(n^3-sum(colSums(x)^3))/(9*n^3)
33 +
34 +     #构建U统计量
35 +     U=z/sqrt(sigma_2)
36 +     if(alternative=="twoside")
37 +     {p_value=1-pchisq(U^2, 1)}
38 +     else
39 +     {
40 +         if(alternative=="greater")
41 +         {p_value=pnorm(-U)}
42 +         else if(alternative=="less")
43 +         {p_value=pnorm(U)}
44 +         else{cat("please input:\n alternative= 'twoside','greater',or'
less'")}
45 +     }
46 +     cat(' 【各种相关系数】 \n')
47 +     cat('Kendall_TAO=',Kendall_TAO,'\n')
48 +     cat('Gamma=',Gamma,'\n')
49 +     cat('d_BA=',d_BA,'\n')
50 +     cat('d_AB=',d_AB,'\n\n')
51 +     cat(' 【相合性检验】 \n')
52 +     cat('U检验统计量的值',U,'\n')
53 +     cat('p_value=',p_value)
54 + }
55
56 > congruence_test(x5)
57 【各种相关系数】
58 Kendall_TAO= 0.4245224
59 Gamma= 0.5719659
60 d_BA= 0.4064291
61 d_AB= 0.4434212
62
63 【相合性检验】
64 U检验统计量的值 8.292197
65 p_value= 1.110223e-16

```

p 值小于  $\alpha=0.01$ , 故拒绝原假设, 认为年龄越大的人, 冠状动脉硬化的程度有越重的趋势.