**附录**

**附录1：**例1的SPSS/R详细操作及结果

**【SPSS软件实现】**：row代表DLN，col代表CLNM，freq代表实际频数。点击[Data]——[Weight Cases],将freq放入Frequency Variable中，点击OK。选择[Analyze]——[Descriptive Statistics]——[Crosstabs],将row放入Row(s)，将col放入Col(s)中，点击Statistics，选择Chi-square，点击OK，结果如下：

附表 1 Chi-Square Tests

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Value | df | Asymp. Sig. (2-sided) | Exact Sig. (2-sided) | Exact Sig. (1-sided) |
| Pearson Chi-Square |  | 1 | 0.024 |  |  |
|  | 3.598 | 1 | 0.058 |  |  |
| Likelihood Ratio | 5.145 | 1 | 0.023 |  |  |
| Fisher’s Exact Test |  |  |  | 0.036 | 0.029 |
| Linear-by-Linear Association | 4.948 | 1 | 0.026 |  |  |
| N of Valid Cases | 41 |  |  |  |  |

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 4.83.

b. Computed only for a 2x2 table

根据软件运行结果附表1，可以看到有1个单元格的理论频数小于5，因此应该采用连续性校正的检验方法，得到=3.598，p=0.058.

**【R软件实现】**：#按行输入数据

example\_1=matrix\_a=matrix(c(20,10,3,8),2,2,byrow=TRUE)

dimnames(example\_1) <- list(“DLN” = c(“positive”, “negative”),

“CLNM” = c(“yes”, “no”))

#卡方连续性校正

chisq.test(example\_1,correct=T)

结果如下：X-squared=3.5983，df=1，p-value=0.05784.

**附录2：**例2的SPSS/R详细操作及结果

**【SPSS软件实现】**：row代表药膏，col代表疗效，freq代表实际频数。点击[Data]——[Weight Cases],将freq放入Frequency Variable中，点击OK。将row放入Row(s)，将col放入Col(s)中，点击Statistics，选择Chi-square，点击OK，结果如下：

附表2 Chi-Square Tests

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Value | df | Asymp. Sig. (2-sided) | Exact Sig. (2-sided) | Exact Sig. (1-sided) |
| Pearson Chi-Square |  | 1 | 0.040 |  |  |
|  | 1.723 | 1 | 0.189 |  |  |
| Likelihood Ratio | 3.262 | 1 | 0.071 |  |  |
| Fisher’s Exact Test |  |  |  | 0.102 | 0.102 |
| Linear-by-Linear Association | 4.124 | 1 | 0.042 |  |  |
| N of Valid Cases | 49 |  |  |  |  |

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 0.61.

b. Computed only for a 2x2 table

根据软件运行结果附表2，可以看到最小的理论频数为0.61，小于1，因此应该采用Fisher确切概率法，得到p=0.102.

**【R软件实现】**：#按行输入数据

example\_2=matrix\_a=matrix(c(38,1,8,2),2,2,byrow=TRUE)

dimnames(example\_2) <- list(“药膏” = c(“药膏A”, “药膏B”),

“疗效” = c(“有效”, “无效”))

#Fisher确切概率

fisher.test（example\_2）

结果如下： p-value=0.1018.

**附录3：**例3的SPSS/R详细操作及结果

**【SPSS软件实现】**：row代表检测方法A，col代表检测方法B，freq代表实际频数。点击[Data]——[Weight Cases],将freq放入Frequency Variable中，点击OK。选择[Analyze]——[Nonparametric Tests]——[Legacy Dialogs]——[2 Related Samples Tests],将row放入Variable1，将col放入Variable2，Test Type选择McNemar,点击OK，结果如下：

附表 3

|  |  |
| --- | --- |
|  | row&col |
| N | 120 |
|  | 11.574 |
| Asymp.Sig. | 0.001 |

a. Continuity Corrected

b. McNemar Test

根据附表3可得：=11.574，p=0.001.

**【R软件实现】**：#输入数据

example\_3=matrix(c(31,14,40,35),2,2,byrow=TRUE)

dimnames(example\_3) <- list(“检测方法A” = c(“阳性”, “阴性”),

“检测方法B” = c(“阳性”, “阴性”))

#McNemar检验，correct=FALSE意味着没有校正

mcnemar.test(matrix\_a,correct=FALSE)

结果如下：McNemar’s chi-squared =12.519,df=1,p-value=0.0004029

可以验证SPSS是进行了连续性校正，mcnemar.test(matrix\_a,correct=TRUE)

结果为：McNemar’s chi-squared =11.574,df=1,p-value=0.0006688，与SPSS结果一致。

**附录4：**例4的SPSS/R详细操作及结果

**【SPSS软件实现】**：row代表静脉血，col代表末梢血，freq代表实际频数。点击[Data]——[Weight Cases],将freq放入Frequency Variable中，点击OK。选择[Analyze]——[Nonparametric Tests]——[Legacy Dialogs]——[2 Related Samples Tests],将row放入Variable1，将col放入Variable2，Test Type选择McNemar,点击OK，结果如下：

附表 4

|  |  |
| --- | --- |
|  | row&col |
| N | 137 |
|  | 3.448 |
| Asymp.Sig. | 0.063 |

a. Continuity Corrected

b. McNemar Test

在SPSS中，根据附表4，=3.448，p=0.063.

**【R软件实现】**：#按行输入数据

example\_4=matrix(c(25,9,20,83),2,2,byrow=TRUE)

dimnames(example\_4) <- list(“静脉血” = c(“阳性”, “阴性”),

“末梢血” = c(“阳性”, “阴性”))

#McNemar检验，correct=TRUE表明进行校正

mcnemar.test(matrix\_a,correct=TRUE)

结果如下：McNemar’s chi-squared =3.4483,df=1,p-value=0.06332.

**附录5：**例5的SPSS/R详细操作及结果

**【SPSS软件实现】**row表示癌症的亚种类型，col表示地区，freq表示实际频数。点击[Data]——[Weight Cases],将freq放入Frequency Variable中，点击OK。选择[Analyze]——[Descriptive Statistics]——[Crosstabs],将row放入Row(s)，将col放入Col(s)中，点击Statistics，选择Chi-square，点击Exact Tests，选择Monte Carlo，点击Continue，最后点击确定。结果如下：

附表 5 Chi-Squard Tests

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Value | df | Asymp.Sig.（2-sided） | Monte Carlo Sig.(2-sided) | | | Monte Carlo Sig.(1-sided) | | |
| Sig. | 99% confidence Intercal | | Sig. | 99% confidence Intercal | |
| Lower Bound | Upper Bound | Lower Bound | Upper Bound |
| Pearson Chi-Square | 10.039a | 4 | 0.40 | .041b | 0.036 | 0.046 |  |  |  |
| Likelihood Ratio | 9.205 | 4 | 0.56 | .076b | 0.069 | 0.083 |  |  |  |
| Fisher’s Exact Test | 8.644 |  |  | .066b | 0.059 | 0.072 |  |  |  |
| Linear-by-Linear Association | .001c | 1 | 0.973 | 1.000b | 1.000 | 1.000 | .527b | .514 | .540 |
| N of Valid Cases | 61 |  |  |  |  |  |  |  |  |

a. 4 cells(44.4%) have expected count less than 5.The minimum expected count is 3.44.

b. Based on 10000 sampled tables with starting seed 2000000.

c. The standardized statistic is -.034.

从结果上可以看出有4 单元格(44.4%) 的期望计数少于 5，超过了总格子数的1/5，因此采用Fisher确切概率法，得到p=0.066.

**【R软件实现】**：#输入数据

example\_5=matrix(c(2,3,10,8,3,5,4,10,16),3,3,byrow=TRUE)

dimnames(example\_5) <- list(“类型” = c(“亚种1”, “亚种2”,”亚种3”),

“地区” = c(“地区1”, “地区2”,”地区3”))

#Fisher确切概率法

fisher.test(matrix\_a,correct=TRUE)

结果如下：p-value=0.06593.

**附录6：**例6的SPSS/R详细操作及结果

**【SPSS软件实现】**row表示药物，col表示疗效，freq表示实际频数。点击[Data]——[Weight Cases],将freq放入Frequency Variable中，点击OK。点击【Analyze】——【Nonparametric Tests】——【Legacy Dialogs】——【k Independent Samples】，将col放入Test Variable List，将row放入Grouping Variable，并Define Range:Minimum为1，Maximum为3，Test Type选择Kruskal-Wallis H，点击OK，结果如下：

附表 6

|  |  |
| --- | --- |
|  | col |
| Chi-Square | 2.830 |
| df | 2 |
| Asymp.sig. | 0.243 |

a. Kruskal-Wallis Test

b. Grouping Variable：col

**【R软件实现】**library(rstatix)

#按行输入数据

example\_6=matrix(c(6,7,20,9,5,9,11,22,10,5,8,15),3,4,byrow=T)

dimnames(example\_6) <- list(“药物” = c(“药物A”, “药物B”, “药物C”),

“疗效” = c(“无效”, “好转”, “显效”,”治愈”))

#将列联表转化为数据框

example\_6\_df <- counts\_to\_cases(example\_6)

example\_6\_df$疗效 <- as.numeric(example\_6\_df$疗效)

example\_6\_df$药物 <- factor(example\_6\_df$药物)

#进行KW检验

kruskal.test(疗效 ~ 药物, example\_6\_df)

结果如下：

Kruskal-Wallis chi-squared = 2.8302, df = 2, p-value = 0.2429.

**附录7：**例7的SPSS/R详细操作及结果

**【SPSS软件实现】**row表示年龄，col表示高血压，freq表示频数。点击[Data]——[Weight Cases],将freq放入Frequency Variable中，点击OK。选择[Analyze]——[Descriptive Statistics]——[Crosstabs],将row放入Row(s)，将col放入Col(s)中，点击Statistics，选择Chi-square，点击OK，结果如下：

附表 7 Chi-square Tests

|  |  |  |  |
| --- | --- | --- | --- |
|  | Value | df | Asymp.sig.（2-sided） |
|
|
| Pearson Chi-Square | 28.444a | 3 | 0.000 |
| Likelihood Ratio | 30.247 | 3 | 0.000 |
| Linear-by-Linear Association | 27.791 | 1 | 0.000 |
| N of Valid Cases | 497 |  |  |

根据附表7，LLAT value为27.791，p<0.001.

**【R软件实现】**#按行输入数据

example\_7=matrix(c(15,85,27,76,75,115,48,56),4,2,byrow=TRUE)

dimnames(example\_7)=list(“年龄”=c(“18岁以下”, “18-35岁”, “35-60岁”,”60岁以上”),

“高血压”=c(“有”,”无”))

#加载包

install.packages(“DescTools”)

library(DescTools)

#进行CATT

CochranArmitageTest(example\_7)

结果：Z = 5.277, dim = 4, p-value = 1.313e-07.

**附录8：**例8的SPSS/R详细操作及结果

**【SPSS操作过程】**row代表检测方法B，col代表检测方法A，freq表示实际频数。点击[Data]——[Weight Cases],将freq放入Frequency Variable中，点击OK。选择[Analyze]——[Descriptive Statistics]——[Crosstabs], 将row放入Row(s)，将col放入Col(s)中，点击Statistics，选择Kappa和McNemar，点击OK。结果如下：

附表 8 Chi-Square Tests

|  |  |  |  |
| --- | --- | --- | --- |
|  | Value | df | Asymp.sig.(2-sided) |
| McNemar-BowkerTest | 8.88 | 3 | 0.031 |
| N of Valid Cases | 103 |  |  |

附表 9 Symmetric Measures

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Value |  | Approx. | Approx.Sig. |
|
|
| Measures of Agreement Kappa | 0.323 | 0.071 | 4.802 | 0.000 |
| N of Valid Cases | 103 |  |  |  |

1. Not assuming the null hypothesis
2. Using the asymptotic standard erroe assuming the null hypothesis

根据实验目的，选择合适的检验方法，如果是检验这两种检测方法是否存在差异，可以根据附表8可以得到=8.88，p=0.031。如果看检测结果是否具有一致性，根据附表9可以得到Kappa值为0.323，p＜0.05。

**【R软件实现过程】**#按行输入数据

example\_8=matrix(c(19,10,8,4,16,3,9,13,21),3,3,byrow=TRUE)

dimnames(example\_8) <- list(“检测方法B” = c(“-“, “+”, “++”),

“检测方法A” = c(“-“, “+”, “++”))

#McNemar-Bowker对称性检验：

mcnemar.test(example\_8,correct=F)

结果如下：McNemar’s chi-squared = 8.8803, df = 3, p-value = 0.03093

#Kappa检验：

Library(irr)

#将列联表转化为数据框

example\_8\_df=counts\_to\_cases(example\_8)

example\_8\_df$检测方法B <- factor(example\_7\_df$检测方法B)

example\_8\_df$检测方法A <- factor(example\_7\_df$检测方法A)

kappa(example\_7\_df)

结果如下：

附表10 Kappa值

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Estimate | Std.Err | 2.5% | 97.5% | p-value |
| kappa | 0.3231 | 0.07096 | 0.184 | 0.4622 | 5.275e-06 |

**附录9：**例9的SPSS/R详细操作及结果

Kruskal-Wallis检验过程与附录7一样, 线性趋势分析与Kendall’s tau-b等级相关分析步骤如下：**【SPSS操作过程】**线性趋势分析：row代表年龄，col代表疗效，freq表示实际频数。点击[Data]——[Weight Cases],将freq放入Frequency Variable中，点击OK。选择[Analyze]——[Descriptive Statistics]——[Crosstabs], 将row放入Row(s)，将col放入Col(s)中，点击Statistics，选择Chi-Square，点击OK。结果如下：

附表 11 Chi-Square Tests

|  |  |  |  |
| --- | --- | --- | --- |
|  | Value | df | Asymp.sig.（2-sided） |
|
|
| Pearson Chi-Square | 6.348a | 4 | 0.175 |
| Likelihood Ratio | 5.945 | 4 | 0.203 |
| Linear-by-Linear Association | 1.459c | 1 | 0.227 |
| N of Valid Cases | 206 |  |  |

1. 0 cells(0%) have expected count less than 5.The minimum expected count is 6.14.

根据附表11中的线性和线性组合，可得=1.459，p=0.227。

Kendall’s tau-b等级相关分析：选择[Analyze]——[Correlate]——[Bivariate Correlations]，将row与col放入Variables中，选择Kendall’s tau-b，点击OK。结果如下表：

附表 12 Correlations

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | row | col |
| Kendall’s tau\_b | row | Correlation Coefficient | 1 | -0.057 |
| Sig.（2-tailed） |  | 0.368 |
| N | 206 | 206 |
| col | Correlation Coefficient | -0.057 | 1 |
| Sig.（2-tailed） | 0.368 |  |
| N | 206 | 206 |

根据附表12可得：相关系数为-0.057，p=0.368.

**【R软件实现过程】**:

example\_9=matrix(c(8,32,45,4,25,37,11,18,26),3,3,byrow=TRUE)

dimnames(example\_9) <- list(“年龄” = c(“18岁以下”, “18-45岁”, “45岁以上”),

“疗效” = c(“无效”, “好转”, “治愈”))

线性趋势分析：

linear\_trend=function(table){

#计算总体的值，可以具体选择连续性校正或者Fisher确切概率法

chisq\_total=chisq.test(table)$statistic

#等距赋值

value\_row=seq(1,nrow(table))

value\_col=seq(1,ncol(table))

n\_row=apply(table,1,sum)

n\_col=apply(table,2,sum)

n\_row\_x=n\_row\*value\_row

n\_row\_x2=n\_row\*value\_row^2

n\_col\_y=n\_col\*value\_col

n\_col\_y2=n\_col\*value\_col^2

#计算

Lxx=sum(n\_row\_x2)-sum(n\_row\_x)^2/sum(n\_row)

#计算

Lyy=sum(n\_col\_y2)-sum(n\_col\_y)^2/sum(n\_col)

#计算

value\_xy=matrix(0,nrow(table),ncol(table))

for(I in 1:nrow(table)){

for (j in 1:ncol(table)){

value\_xy[i,j]=table[i,j]\*value\_row[i]\*value\_col[j]

}

}

#计算

Lxy=sum(apply(value\_xy,1,sum))-sum(n\_row\_x)\*sum(n\_col\_y)/sum(n\_row)

#计算线性回归分量和对应的p值

chisq\_reg=sum(ni)\*Lxy^2/Lxx/Lyy

reg\_p\_value=pchisq(chisq\_reg,1,lower.tail = FALSE)

#计算偏离线性回归分量和对应的p值

chisq\_dev=chisq\_total -chisq\_reg

dev\_p\_value=pchisq(chisq\_dev,(nrow(table)-1)\*(ncol(table)-1)-1,lower.tail = FALSE)

names(chisq\_dev)="chisq\_dev"

names(chisq\_reg)="chisq\_reg"

names(reg\_p\_value)="reg\_p\_value"

names(dev\_p\_value)="dev\_p\_value"

return(list=c(chisq\_reg,chisq\_dev,reg\_p\_value,dev\_p\_value)) }

#代入函数

linear\_trend(example\_9)

结果如下：chisq\_reg chisq\_dev reg\_p\_value dev\_p\_value

1.46568 4.88260 0.2260285 0.1805983

Kendall's tau-b等级相关分析:

#将列联表转化为数据框

library(Kendall)

example\_9\_df <- counts\_to\_cases(example\_9)

example\_9\_df$年龄 <- as.numeric(example\_9\_df$年龄)

example\_9\_df$疗效 <- as.numeric(example\_9\_df$疗效)

# kendall等级相关分析

cor.test(example\_9\_df$年龄,example\_9\_df$疗效,method="kendall")

结果如下:tau=-0.05715426,p-value=0.3682.

**附录11：**例11的SPSS/R详细操作及结果

**【SPSS实现过程】**将数据导入SPSS中，点击[Data]——[Weight Cases]，将freq放入Frequency Variable中，点击OK。选择[Analyze]——[Descriptive Statistics]——[Crosstabs],将分组放入Row（s）中，将高血压放入Col(s)中，将年龄放入Layer中。点击Statistics，选择Cochran’s and Mantel-Haenszel统计量。结果如下：

附表 13 Tests of Homogeneity of the Odds Ratio

|  |  |  |  |
| --- | --- | --- | --- |
|  | Chi-Squared | df | Asymp.sig.(2-sided) |
| Breslow-Day | 2.591 | 1 | 0.107 |
| Tarone’s | 2.591 | 1 | 0.107 |

附表 14 Tests of Conditional Independence

|  |  |  |  |
| --- | --- | --- | --- |
|  | Chi-Squared | df | Asymp.sig.(2-sided) |
| Cochran’s | 31.013 | 1 | 0.000 |
| Mantel-Haenszel | 30.272 | 1 | 0.000 |

**【R软件实现】**

library(DescTools)

example\_11 <- array(c(155,53,

88,43,

213,165,

145,254),

dim = c(2,2,2),

dimnames = list(心肌梗塞 = c("病例","对照"),

高血压 = c("是","否"),

年龄分层 = c("<35岁","≥35岁")

)

)

#Breslow-Day检验

BreslowDayTest(example\_11)

结果如下：X-squared = 2.5907, df = 1, p-value = 0.1075

#分层检验

mantelhaen.test(example\_11,correct = F)

结果如下：Mantel-Haenszel X-squared = 30.96, df = 1, p-value =2.634e-08