

▼ 三軍總醫院北投分院統計及實驗設計課程之六

2021/7/7

ytai1123@gmail.com

使用方法:

1. 使用gmail帳號登入
2. 按"執行階段" --> "全部執行" 以執行全部內容, 若要個別執行可點選每格程式左方箭頭或按 Control + Enter 鍵執行。

##0-1

```
!git clone https://github.com/YuehMintTai/RPython.git
```

```
Cloning into 'RPython'...
remote: Enumerating objects: 213, done.
remote: Counting objects: 100% (213/213), done.
remote: Compressing objects: 100% (211/211), done.
remote: Total 213 (delta 120), reused 0 (delta 0), pack-reused 0
Receiving objects: 100% (213/213), 1005.97 KiB | 6.06 MiB/s, done.
Resolving deltas: 100% (120/120), done.
```

##0-2

```
!pip install rpy2
```

```
Requirement already satisfied: rpy2 in /usr/local/lib/python3.7/dist-packages (3.4.5)
Requirement already satisfied: pytz in /usr/local/lib/python3.7/dist-packages (from rpy2) (2020.5)
Requirement already satisfied: Jinja2 in /usr/local/lib/python3.7/dist-packages (from rpy2) (2.11.3)
Requirement already satisfied: tzlocal in /usr/local/lib/python3.7/dist-packages (from rpy2) (2.1)
Requirement already satisfied: cffi>=1.10.0 in /usr/local/lib/python3.7/dist-packages (from rpy2) (1.14.6)
Requirement already satisfied: MarkupSafe>=0.23 in /usr/local/lib/python3.7/dist-packages (from Jinja2->rpy2) (2.0.1)
Requirement already satisfied: pycparser in /usr/local/lib/python3.7/dist-packages (from cffi->rpy2) (2.21)
```

##0-3

```
%load_ext rpy2.ipynb
```

##6-1

```
%%R
```

```
myData<-read.csv('RPython/samples.csv')
```

```
tail(myData, 1)
```

```

SID 性別 年齡 入伍前職業 教育程度 婚姻狀況 皆無過去病史01 早產兒01
188  4    1   25      商      4      1      1      0
    頭部曾受傷01 發展遲緩01 注意力不足過動症01 癲癇01 癲癇服藥治療 癲癇服藥期間
188      0      0      0      0      0      0
    軍種 軍階 役別 入伍至今_年 聽過自殺課程_次 求助心輔_次 求助精神科_次

```

```

188      1      1      2      0.5      1      0      2
      使用1995_次 使用24h專線_次 特殊狀況 父母婚姻狀態 自殺意念_bsrs6 B型肝炎01
188      0      0      4      4      4      0
      C型肝炎01 氣喘史01 過敏史01 心臟病史01 高血壓01 糖尿病01 甲狀腺01 類風濕01
188      0      1      1      0      0      0      1      0
      重大意外01 自殺意念01 透露父母 透露手足 透露好友 透露同儕 透露長官 透露心輔
188      1      1      0      0      0      0      0      0
      透露醫師 拒告父母 拒告手足 拒告好友 拒告同儕 拒告長官 拒告心輔 拒告醫師
188      0      1      1      1      1      1      1      1
      BSRS總分 BSRSR總分 過動症總分 Inattention Impulsivity opposition depression
188      20      5      18      9      9      8      57
      anxiety burdensome belonging 家庭滿意度apgar 網路成癮症01 網路成癮分數YDQ
188 29.0294      42      12      0      0      0
      existenness meaning control seeking death suicidea 睡眠困擾_bsrs1
188      28      10      22      16      15      7      4
      睡眠困擾_bsrsr1 睡眠困擾_bdi16 易怒_bsrs3 易怒_bsrsr3 depress impuls
188      1      3      4      1      57      9
      Internet ADHD
188      0      18

```

##6-2

%%R

formula<-'自殺意念01~as.factor(性別)'

modell<-glm(formula, myData,family='binomial')

exp(cbind(OR=coef(modell), confint(modell)))

R[write to console]: Waiting for profiling to be done...

```

              OR      2.5 %      97.5 %
(Intercept)    0.2384615 0.1584208 0.3477816
as.factor(性別)2 0.7293128 0.2030315 2.0664801

```

##6-3

%%R

myData\$自殺意念01<-as.integer(myData\$自殺意念01)

myTable<-table(myData\$自殺意念01, myData\$性別)

rownames(myTable)<-c('0無自殺意念','1有自殺意念')

colnames(myTable)<-c('1男性','2女性')

table2x2(myTable, stats=c('table','or'))

2x2 contingency table

	1男性	2女性	Sum
0無自殺意念	130	23	153
1有自殺意念	31	4	35
--	--	--	--
Sum	161	27	188

Statistics

a= 130

```
b= 23
c= 31
d= 4
```

```
p1=a/(a+b)= 0.8497
p2=c/(c+d)= 0.8857
```

Odds ratio

```
Odds ratio = OR = (p1/(1-p1))/(p2/(1-p2)) = 0.7293
Standard error = SE.OR = sqrt((1/a+1/b+1/c+1/d)) = 0.5774
Lower 95%-confidence limit: = OR * exp(- 1.96 * SE.OR) = 0.2352
Upper 95%-confidence limit: = OR * exp(1.96 * SE.OR) = 2.2617
```

The estimated odds ratio is 0.729 (CI_95%: [0.235;2.262]).

```
##6-3-1
```

```
%%R
```

```
with(myData, table(自殺意念01, 性別))
```

```
      性別
自殺意念01  1   2
0      130  23
1       31   4
```

```
##6-4
```

```
import pandas as pd
import numpy as np
import statsmodels.formula.api as smf
formula='自殺意念01.astype(int)~C(性別)'
model3=smf.logit(formula=formula,data=df)
result=model3.fit()
result.summary()
print('OR={}, 95%CI={} to {}'.format(round(np.exp(result.params)[1],3),\
round(np.exp(result.conf_int())[0][1],3),\
round(np.exp(result.conf_int())[1][1],3))
```

Optimization terminated successfully.

Current function value: 0.479782

Iterations 6

OR=0.729, 95%CI=0.235 to 2.262

```
##6-4
```

```
import pandas as pd
import scipy.stats as stats
df=pd.read_csv('RPython/samples.csv')
df['自殺意念01']=df['自殺意念01'].astype(int)
myTable=pd.crosstab(df['自殺意念01'],df['性別'])
print(myTable)
(oddsratio, pvalue)=stats.fisher_exact(myTable, alternative='two-sided')
print('Odds Ratio:{}, p-value:{}'.format(round(oddsratio,2),round(pvalue,3)))
```

性別	1	2
自殺意念01		
0	130	23
1	31	4

Odds Ratio:0.73, p-value:0.79

```
##6-5-1 MultiNomial logistic regression
%%R
with(myData, table(自殺意念_bsrs6, 性別))
```

	性別	
自殺意念_bsrs6	1	2
0	118	21
1	25	5
2	9	1
3	7	0
4	2	0

```
%%R
install.packages('multinom')
```

R[write to console]: Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

```
%%R
library(multinom)
formula<- '自殺意念_bsrs6~年齢+as.factor(性別)'
result<-multinom(formula, myData)
```



R[write to console]: Error in library(multinom) : there is no package called 'multinom'

R[write to console]: In addition:

R[write to console]: Warning message:

R[write to console]: package 'multinom' is not available for this version of R

A version of this package for your version of R might be available elsewhere, see the ideas at

<https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-packages>

Error in library(multinom) : there is no package called 'multinom'

```
-----
RRuntimeError                                Traceback (most recent call last)
/usr/local/lib/python3.7/dist-packages/rpy2/ipython/rmagic.py in eval(self, code)
    267                                     # Need the newline in case the last line in code is a
comment.
--> 268                                     value, visible = ro.r("withVisible({%s\n})" % code)
    269                                     except (ri.embedded.RRuntimeError, ValueError) as exception:
```

%%R

```
##install.packages('Publish')
```

```
remove.packages('Publish')
```

R[write to console]: Removing package from '/usr/local/lib/R/site-library'
(as 'lib' is unspecified)

```
<decorator-gen-119> in R(self, line, cell, local_ns)
```

%%R

```
library(Publish)
```

```
myTable<-data.frame(Case=c(12,34),Control=c(56,78))
```

```
rownames(myTable)<-c('Exposure','No exposure')
```

```
table2x2(myTable,stats=c('table','or','fisher'))
```

2x2 contingency table

	Case	Control	Sum
Exposure	12	56	68
No exposure	34	78	112
--	--	--	--
Sum	46	134	180

Statistics

a= 12

b= 56

c= 34

d= 78

p1=a/(a+b)= 0.1765

p2=c/(c+d)= 0.3036

Odds ratio

Odds ratio = OR = $(p1/(1-p1))/(p2/(1-p2)) = 0.4916$
 Standard error = SE.OR = $\sqrt{(1/a+1/b+1/c+1/d)} = 0.3787$
 Lower 95%-confidence limit: = $OR * \exp(-1.96 * SE.OR) = 0.2340$
 Upper 95%-confidence limit: = $OR * \exp(1.96 * SE.OR) = 1.0327$

The estimated odds ratio is 0.492 (CI_95%: [0.234;1.033]).

Fisher's exact test

Fisher's Exact Test for Count Data

```
data:  table2x2
p-value = 0.07749
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.2129831 1.0818070
sample estimates:
odds ratio
 0.4934583
```

```
%%R
```

```
table2x2(myTable, stats=c('table','or','fisher'))
```

2x2 contingency table

	Case	Control	Sum
Exposure	12	56	68
No exposure	34	78	112
--	--	--	--
Sum	46	134	180

Statistics

a= 12
 b= 56
 c= 34
 d= 78

$p1=a/(a+b) = 0.1765$
 $p2=c/(c+d) = 0.3036$

Odds ratio

Odds ratio = OR = $(p1/(1-p1))/(p2/(1-p2)) = 0.4916$

Standard error = SE.OR = $\sqrt{(1/a+1/b+1/c+1/d)} = 0.3787$

Lower 95%-confidence limit: = $OR * \exp(-1.96 * SE.OR) = 0.2340$

Upper 95%-confidence limit: = $OR * \exp(1.96 * SE.OR) = 1.0327$

The estimated odds ratio is 0.492 (CI_95%: [0.234;1.033]).

Fisher's exact test

Fisher's Exact Test for Count Data

data: table2x2

p-value = 0.07749

alternative hypothesis: true odds ratio is not equal to 1

95 percent confidence interval:

0.2129831 1.0818070

sample estimates:

odds ratio

0.4934583