

類別資料推論(一)

Inference for Categorical Data(1)

嚴明芳 助理教授
台北醫學大學口腔衛生學系
2011/11/22

生物統計學
Biostatistics

辦公室：口腔醫學大樓4樓
Email : amyen@tmu.edu.tw
分機：5152

類別資料分析
(Categorical Data Analysis)

學習目標及資料來源

- 學習目標

- 使學生了解類別資料推論之相關統計概念

- 資料來源

- [沈明來](#)：生物統計學入門 5/e (2007)
ISBN：9789866929021，九州出版社。

- 吸菸、戒菸和不吸菸的比例和居住地無關？

吸菸習慣	信義區	文山區	中山區	總計
從不	2146	110	2293	4549 (78.2%)
已戒	182	12	185	379 (6.5%)
目前有吸	402	28	458	888 (15.3%)
總計	2730	150	2936	5816

獨立性檢定—卡方檢定

- 觀察值 (Observed value) O
- 期望值 (Expected value) E
- 卡方值 (Chi-square)

$$\sum \frac{(O - E)^2}{E} \quad \text{cf. Chi-square distribution (df=?)}$$

推論

- 吸菸、戒菸和不吸菸的比例和居住地無關？

吸菸習慣	信義區	文山區	中正區	總計
從不				4549 (78.2%)
已戒				379 (6.5%)
目前有吸				888 (15.3%)
總計	2730	150	2936	5816

觀察值

吸菸習慣	信義區	文山區	中正區	總計
從不	2146	110	2293	4549 (78.2%)
已戒	182	12	185	379 (6.5%)
目前有吸	402	28	458	888 (15.3%)
總計	2730	150	2936	5816

Chi-sq=3.20

期望值

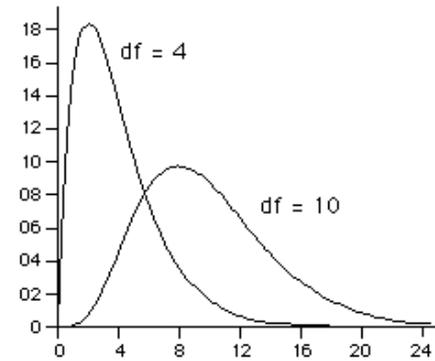
吸菸習慣	信義區	文山區	中正區	總計
從不	2135.28	117.32	2296.40	4549 (78.2%)
已戒	177.90	9.77	191.32	379 (6.5%)
目前有吸	416.82	22.90	448.28	888 (15.3%)

2011/11/22

總計 2730 150 2936 5816

口腔衛生學系 生物統計學

卡方分佈



- 已知卡方值及自由度，如何求算p值...
 - 專業統計軟體
 - Excel

Microsoft Excel - Book2

檔案(F) 編輯(E) 檢視(V) 插入(I) 格式(O) 工具(T) 資料(D) 視窗(W)

新細明體 12 B I U \$ % , +.00

回覆變更(C)... 結束檢閱(N)...

	A	B	C	D	E	F
1	Chisq	d.f	P-value			
2	3.2	4	0.524931			
3						
4	口腔衛生學系	生物統計學				7
5						
6						

Fisher Exact Test

Type of diet	Cause of death		Total
	Non-CVD	CVD	
High salt	2	5	7
Low salt	23	30	53
Total	25	35	60

```

data a;
  input Diet COD n;
cards;
  1 1 2
  1 2 5
  2 1 23
  2 2 30
;
proc freq;
  table Diet*COD/chisq;
  weight n;
run;

```

Statistics for Table of C by D

Statistic	DF	Value	Prob
Chi-Square	1	0.5591	0.4546
Likelihood Ratio Chi-Square	1	0.5811	0.4459
Continuity Adj. Chi-Square	1	0.1155	0.7339
Mantel-Haenszel Chi-Square	1	0.5498	0.4584
Phi Coefficient		-0.0965	
Contingency Coefficient		0.0961	
Cramer's V		-0.0965	

WARNING: 50% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Fisher's Exact Test

Cell (1,1) Frequency (F)	2
Left-sided Pr <= F	0.3747
Right-sided Pr >= F	0.8775
Table Probability (P)	0.2522
Two-sided Pr <= P	0.6882

Sample Size = 60

Exact probability of observing a table as follows

Type of diet	Cause of death		Total
	Non-CVD	CVD	
High salt	a	b	a+b
Low salt	c	d	c+d
Total	a+c	b+d	n

$$\Pr(a, b, c, d) = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!}$$

For the previous example

0	7
25	28

.017

1	6
24	29

.105

2	5
23	30

.252

3	4
22	31

.312

4	3
21	32

.214

5	2
20	33

.082

6	1
19	34

.016

7	0
18	35

.001

McNemar's檢定的應用

- 流行病學
 - 在流行病學中病例對照研究法

McNemar's 檢定

		後測	
		答對	答錯
前測	答對	a	b
	答錯	c	d

Yate's correction

$$\chi^2 = \frac{\left(b - \frac{n_D}{2}\right)^2}{\frac{n_D}{4}}$$
$$\chi^2 = \frac{\left(\left|b - \frac{n_D}{2}\right| - \frac{1}{2}\right)^2}{\frac{n_D}{4}}$$

2011/11/22 若 $X^2 > \chi^2_{1,1-\alpha}$ 則拒絕虛無假說 $H_0: p=1/2$

McNemar's 檢定

- $n_D < 20 \rightarrow$ Exact test

$$p = 2 \times \sum_{k=0}^{n_A} \binom{n_D}{k} \left(\frac{1}{2}\right)^{n_D} \quad \text{if } n_A < n_D/2$$

$$p = 2 \times \sum_{k=n_A}^{n_D} \binom{n_D}{k} \left(\frac{1}{2}\right)^{n_D} \quad \text{if } n_A \geq n_D/2$$

一致性檢定

6.請問您是否有每月量體重的習慣？0.不曾 1.有

		第二位訪視者		小計
		不曾	有	
第一位訪視者	不曾	160	60	220
	有	40	240	280
小計		200	300	500

一致性檢定

- Reliability Index (RI)：兩次測量相同比例， $(a+d) / N$
- $RI = (160 + 240) / 500 = 80\%$
- $\kappa = 0.5902$

Kappa係數

$$K = \frac{P_O - P_E}{1 - P_E}$$

- $P_O = RI$

$$P_E = \frac{\left\{ [(a + b) \times (a + c) / N] + [(c + d) \times (b + d) / N] \right\}}{N}$$

- Kappa值 ≤ 0.40 代表一致性不好，
- Kappa值介於 $0.41 \sim 0.60$ 代表一致性普通，
- Kappa值介於 $0.61 \sim 0.80$ 代表有不錯的一致性，
- Kappa值介於 $0.81 \sim 1.00$ 代表有極高的一致性。

$$se(k) = \sqrt{\frac{1}{N(1-P_e)^2} \times \left\{ P_e + P_e^2 - \sum_{i=1}^c [a_i b_i (a_i + b_i)] \right\}}$$

欲檢定虛無假說 $H_0: \kappa=0$ ，其對立假說為 $H_1: \kappa > 0$ ，其檢定統計量為

$$z = \frac{\kappa}{se(\kappa)} \sim N(0,1)$$

若 $z > z_{1-\alpha}$ 拒絕虛無假說，否則則接受虛無假說

$$P = 1 - \Phi(z)$$

範例

- 訪視口服避孕藥服用史所得2×2 Table 如下

		第二位訪視者		小計
		不曾	有	
第一位訪視者	不曾	60	10	70
	有	10	20	30
小計		70	30	100

RI=?

2011/11/2 K=?

總結

- 獨立性檢定—卡方檢定
- Fisher Exact Test
- McNemar's檢定及應用
- 一致性檢定