類別變項的統計方法（續）

林明瀅
台北榮民總醫院感染管制委員會

叁、費歇恰當檢定（Fisher's exact test）

其原理相當簡單，即將直行與橫列的和當做常數或固定值，正確的（exactly）來計算各細格（cell）之排列組合會達到所觀察到的那種極端或更極端的機率。

一、適用條件：1. 2x2卡方檢定，若N＜20且其中超過20％的細格小於5時。

二、假設檢定：有無曝露於某危險因子的得病機率相同

三、檢定程序：

1) 檢定每一細格之期望值是否過小（＜5）？

2) 檢定有無曝露於某危險因子的得病機率是否相同？

四、檢定統計量：

a	b	a+b
c	d	c+d
a+c	b+d	N

a,b,c,d為4個細格的數值
N為總樣本數

$$Pi=\frac{(a+b)!(c+d)!(a+c)!(b+d)!}{N!a!b!c!d!}$$

五、範例：某研究者研究17位呼吸窘迫的患者，其使用吸呼器與肺炎的關係，其資料如下：

（設α=0.05）

		使用呼吸器		合　計
		是	否	
得肺炎	是	4	1	5
	否	6	6	12
合　計		10	7	17

計算過程：在計算時，先找數目最小之細格，計算其Pi值，再將最小細格減1，重新調整其他三個細格的數目，（亦即分佈更為極端），再計算Pi，反覆為之，直至最小細格數為0為止。將各Pi值相加，即為P值。

1)Pi即為該特定排列組合之機率

$$Pi=\frac{5!\times 12!\times 7!\times 10!}{17!\times 1!\times 4!\times 6!\times 6!}=0.238$$

$$(*5!=5\times 4\times 3\times 2\times 1)$$

2)比上述更極端的排列組合則為

5	0	5
5	7	12
10	7	17

此特定排列組合發生之機率則為

$$Pi=\frac{5!\times 12!\times 7!\times 10!}{17!\times 0!\times 5!\times 7!\times 5!}=0.041$$

(*0!=1)

3) P=.238+.041=.279>0.05

六、檢定結果：P>0.05，則接受無差異之假設，即使用呼吸器與肺炎無關。

七、決策原則：

1. 檢測預期值是否小於或等於5，最簡單的試探方法是從行與列之邊緣總數（marginal totals）中各選較小者，然後再兩者相乘，再除以（總數），所得的值是否小於或等於5。

2. 預期值小於或等於5，不宜用一般的卡方檢定其理由有二：一為樣本數可能過小，因此無法使每個細格的預期值大於5。第二種情形雖然樣本數不小，但p（某事件發生的機率，如發生院內感染），或（1－p）（某件事不發生的機率，如不發生院內感染），兩者之一太小，如p很小，則（a+c）×p或（b+d）×p就不易達到5以上。因預期值很小而達不到5，則因機會引起差別的可能性增大（不是真正的差別機會增大）。

肆、McNemar's test for paired sample

McNemar氏考驗又稱為非獨立樣本比率數的卡方考驗（Chi－square test for independent sample proportions），或稱為相依樣本的卡方檢定。主要探討實驗前後由『是』變為『否』與由『否』變為『是』的個數是否相等，因此被稱為『改變的顯著檢定』。

一、適用條件：

1.類別變項—類別變項

2.兩相依樣本（即相同人，在不同時間做測量或配對資料）

3.僅適用於2x2的表格

4.樣本是從母群體中隨機抽樣

二、假設檢定：實驗前後，選擇『是』的百分比沒有改變

三、檢定程序：檢定實驗前後，選擇『是』的百分比是否沒有改變？

四、檢定統計量：此時直行與橫行不再分別代表2個變數，而是表示同一變數在不同時間（情況）下，先後『重複測驗』的結果。其資料形式可排成如下格式：

		實驗後（或實驗組）		
		是	否	
實驗前 控制組	是	A	B	A+B
	否	C	D	C+D
		A+C	B+D	N

A,B,C,D為4個細格的數值

N為總樣本數

$$X^2=\frac{(|B-C|-1)^2}{B+C}$$

H_0:P1＝P2或B＝C

H_1:P1≠P2或B≠C

五、範例：某研究者想了解80名新進人員職前教育，對肺結核患者是否需採取『耐酸性桿菌隔離』的認知，採用在授課前後兩次測量其資料如下

		上課後 是	上課後 否	
上課前	是	11	2	13
上課前	否	47	20	67
		58	22	80

計算過程（代入公式）

$$X^2 = \frac{(|47-2|-1)^2}{47+2} = 39.51$$

六、檢定結果：$39.51 > X^2 99(1) = 6.63 \Rightarrow P$,$P < 0.01$　達統計顯著拒絕H_0，即職前教育改變了新進人員的隔離技術觀念。

七、決策原則

如果是重複測量，總樣本數N，即代表擁有N個人（樣本）參與研究。若是配對資料（即實驗組與控制組配對），則總樣本數N代表有N樣本參與研究，實際上是2N個人。

伍、波以松檢定（Poisson test）

一、適用條件：某事件，發生的機率很小，亦即p值很小，通常在0.01以下，而q很大，通常在0.99以上，而且此機率為已知或恆定，因此即使抽取相當大的樣本數，$n \times p$仍然小於5。

二、假設檢定：兩組資料之發生率相等。

三、檢定程序：檢定兩組資料之發生率是否相等？

四、檢定統計量：

波以松分佈的公式　$P(x) = \dfrac{e^{-\lambda} \times \lambda^a}{X!}$

P(a)：為樣本中，某事件正好發生a次的機率

e：為自然對數，即2.7183

λ：為樣本中之預期發生數，即n x p。此處之p亦為某事件之慣常發生率或已知母全體之發生率。

a：為樣本中之實際發生數，亦即觀察值。

五、範例：假設台灣地區近數年來，地區醫院的院內感染發生率為千分之7，若甲醫院79年度住院總人數達1000次，而只有二例院內感染，則是否可宣稱此醫院的院內感染率比全國低？

$H_0:P_1 = P_0$，$H_1:P_1 < P_0$

設$\alpha = 0.05$

計算過程：先計算0次的機率（代入公式），再計算1次的機率，再計算Pi，反覆為之，直至實際發生次數。將各Pi值相加，即為P值。

$\lambda = n \times p = 1000 \times 0.007 = 7$

$$P(0) = \frac{e^{-7} \times 7^0}{0!} = 0.0009$$

$$P(1) = \frac{e^{-7} \times 7}{1!} = 0.0064$$

$$P(2) = \frac{e^{-7} \times 7^2}{2!} = 0.0223$$

$*0! = 1$，$1! = 1$，$2! = 2 \times 1$

發生2次（包括2次）以下之機率為

$0.0009 + 0.0064 + 0.0223 = 0.0296$

$P < \alpha$，甲醫院之院內感染率較全國為低。

*本文範例之數字為假設資料。