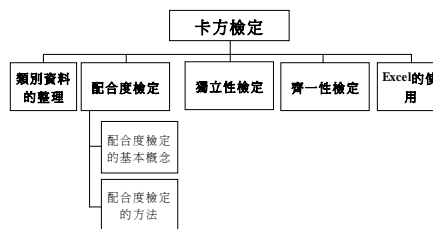


15 類別資料的分析

○ 學習目的

1. 瞭解類別資料的意義與性質。
2. 瞭解當碰到類別資料時，應如何來做統計檢定的工作。
3. 瞭解配合度檢定的意義與檢定方法。
4. 瞭解獨立性檢定的意義與檢定方法。
5. 瞭解齊一性檢定的意義與方法。
6. 利用 Excel 來做卡方檢定。

本章結構



類別資料的分析

○ 類別資料的假設檢定

對於這種類別資料的假設檢定是檢定它的類別比例，檢定的方法是利用卡方分配 (Chi-square distribution) 來做卡方檢定 (Chi-square test)。

類別資料的分析

○ 卡方檢定

卡方檢定主要用來檢定名目尺度資料與順序尺度資料。卡方分配可做三種檢定：配合度檢定、獨立性檢定與齊一性檢定。

類別資料的整理

○ 類別資料的整理方法

類別資料是只能以類別來區分的資料，如男女性別、教育程度、職業別、區域別、滿意程度、偏好程度、品質好壞程度、離職意願等。在處理類別資料時，通常是將樣本依其類別計算其次數，而得到各類別的次數分配表，然後用它來分析資料的特性。

類別資料的整理

表15.1 人民對海峽兩岸關係看法的調查

兩岸關係的看法	人數
維持現狀	1,200
獨立	600
統一	200

類別資料的整理

表15.2 列聯表-性別與喜好車型

性別\車型	中型車	小型車	迷你車
男	30	25	25
女	15	20	35

配合度檢定

○ 配合度檢定的意義

配合度檢定是利用樣本資料檢定母體分配是否為某一特定分配或理論分配的統計方法。

○ 多項實驗

- ① 包含 n 次相同的試行。
- ② 每一次試行結果為 k 個可能結果中的一個($k > 2$)。
- ③ 每一次試行是獨立的。
- ④ 每一次試行各個可能結果的機率是固定的。即若每一結果的機率為 $p_i (i = 1, \dots, k)$ ，則每一次試行結果的機率均相同且 $p_1 + p_2 + \dots + p_k = 1$ 。

配合度檢定

○ 觀察次數與期望次數

抽樣實驗中得到的次數稱為觀察次數。表為 O ；期望次數是如果虛無假設為真時，預期會發生的次數，表為 E 。各類別的期望次數為：

$$E = np$$

n 為樣本數， p 為虛無假設為真時，各類別的比例或機率。

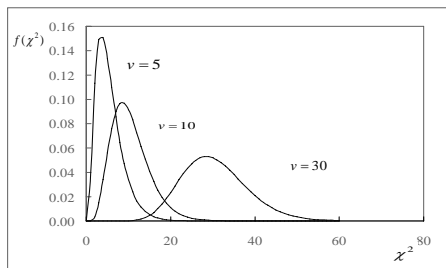
配合度檢定

○ 卡方分配的特性

- ① 卡方分配為一定義在大於等於0（正數）範圍的右偏分配，不同的自由度決定不同的卡方分配。
- ② 卡方分配只有一個參數即自由度，表為 ν 。卡方分配的平均數 $E(\chi^2) = \nu$ 為變異數為 $V(\chi^2) = 2\nu$ 。
- ③ 卡方分配隨自由度的增加而逐漸對稱，當自由度趨近於無窮大時($\nu \rightarrow \infty$)，卡方分配會趨近於常態分配。
- ④ 設 $X \sim N(\mu, \sigma^2)$ ，令 $Z^2 = (\frac{X-\mu}{\sigma})^2$ ，則 Z^2 為自由度1的卡方分配。

配合度檢定

圖 15.1 卡方分配



配合度檢定

表15.3 卡方值

df	$\chi^2_{0.995}$	$\chi^2_{0.95}$	$\chi^2_{0.90}$	$\chi^2_{0.80}$	$\chi^2_{0.70}$	$\chi^2_{0.60}$	$\chi^2_{0.50}$	$\chi^2_{0.40}$	$\chi^2_{0.30}$	$\chi^2_{0.20}$	$\chi^2_{0.10}$	$\chi^2_{0.05}$	$\chi^2_{0.025}$	$\chi^2_{0.01}$	$\chi^2_{0.005}$	df
1	0.000193	0.003932	0.0157908	2.70554	3.84146	5.02389	6.63489	7.87944	9.00079	10.5966	12.0008	13.2767	14.4494	15.4919	16.2660	1
2	0.0100251	0.02587	0.102587	2.10720	4.60517	5.99147	7.37779	8.55759	9.59015	10.5966	11.5913	12.5916	13.5802	14.5379	15.4919	2
3	0.0717212	0.351846	0.584375	1.21276	2.36597	3.21913	4.10846	4.87839	5.54819	6.25139	6.90264	7.37779	7.87944	8.38513	8.90625	3
4	0.206990	0.710721	1.063623	1.75355	2.76579	3.35737	4.07476	4.71317	5.27951	5.77489	6.25139	6.75139	7.10841	7.47989	7.87944	4
5	0.411740	1.145476	1.610131	2.33542	3.35737	4.07476	4.71317	5.27951	5.77489	6.25139	6.75139	7.10841	7.47989	7.87944	5	
6	0.675727	1.63539	2.20413	3.00071	4.07476	4.71317	5.27951	5.77489	6.25139	6.75139	7.10841	7.47989	7.87944	8.23572	6	
7	0.989265	2.16749	2.83306	3.59949	4.60517	5.27951	5.77489	6.25139	6.75139	7.10841	7.47989	7.87944	8.23572	8.53768	7	
8	1.31217	2.73326	3.51724	4.24645	5.27951	5.77489	6.25139	6.75139	7.10841	7.47989	7.87944	8.23572	8.53768	8.90625	8	
9	1.63539	3.33714	4.16832	4.89816	5.77489	6.25139	6.75139	7.10841	7.47989	7.87944	8.23572	8.53768	8.90625	9.23635	9	
10	1.93482	3.93702	4.86813	5.57856	6.25139	6.75139	7.10841	7.47989	7.87944	8.23572	8.53768	8.90625	9.23635	9.59015	10	
11	2.21518	4.57555	5.57856	6.25139	6.75139	7.10841	7.47989	7.87944	8.23572	8.53768	8.90625	9.23635	9.59015	9.90484	11	
12	2.47243	5.22964	6.25139	6.75139	7.10841	7.47989	7.87944	8.23572	8.53768	8.90625	9.23635	9.59015	9.90484	10.2157	12	
13	2.70554	5.89147	6.75139	7.10841	7.47989	7.87944	8.23572	8.53768	8.90625	9.23635	9.59015	9.90484	10.2157	10.5196	13	
14	2.92157	6.55139	7.10841	7.47989	7.87944	8.23572	8.53768	8.90625	9.23635	9.59015	9.90484	10.2157	10.5196	10.8177	14	
15	3.11972	7.14151	7.47989	7.87944	8.23572	8.53768	8.90625	9.23635	9.59015	9.90484	10.2157	10.5196	10.8177	11.1149	15	
16	3.29012	7.69256	7.87944	8.23572	8.53768	8.90625	9.23635	9.59015	9.90484	10.2157	10.5196	10.8177	11.1149	11.4093	16	
17	3.4328	8.20272	8.23572	8.53768	8.90625	9.23635	9.59015	9.90484	10.2157	10.5196	10.8177	11.1149	11.4093	11.6985	17	
18	3.55154	8.67176	8.53768	8.90625	9.23635	9.59015	9.90484	10.2157	10.5196	10.8177	11.1149	11.4093	11.6985	11.9831	18	
19	3.64786	9.11658	8.90625	9.23635	9.59015	9.90484	10.2157	10.5196	10.8177	11.1149	11.4093	11.6985	11.9831	12.2642	19	
20	3.71822	9.54081	9.23635	9.59015	9.90484	10.2157	10.5196	10.8177	11.1149	11.4093	11.6985	11.9831	12.2642	12.5418	20	

配合度檢定

圖15.2 卡方值的對話方塊



配合度檢定

○ 配合度檢定統計量

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \sim \chi_{k-1-m}^2$$

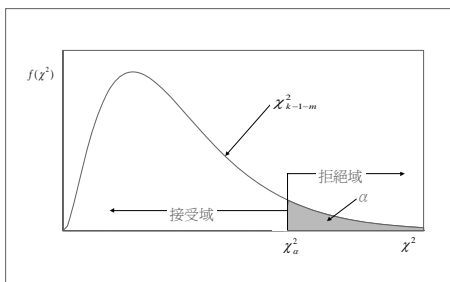
式中： O_i ：第*i*組的樣本觀察次數， E_i ：期望次數(或理論次數)， k ：組數。自由度為 $k-1-m$ ， m 為估計的參數個數。

○ 決策法則

- ① 若 $\chi^2 > \chi_{k-1-m, \alpha}^2$ ，則拒絕 H_0 。
- ② 若 $\chi^2 \leq \chi_{k-1-m, \alpha}^2$ ，則接受 H_0 。

配合度檢定

圖15.3 卡方檢定的拒絕域與接受域



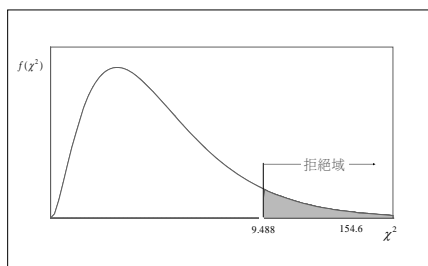
配合度檢定

表15.4 旅客對高雄國際航空站的滿意度

旅客對高雄國際航空站滿意度	人數
非常滿意	157
滿意	463
尚可	118
不滿意	20
非常不滿意	43

配合度檢定

圖15.4 旅客對高雄航空站滿意度的配合度檢定



配合度檢定

表15.5 旅客滿意程度的觀察次數與期望次數

滿意度	O_i	機率 p_i	$E = np_i$	$(O - E)$	$(O - E)^2$	$\frac{(O - E)^2}{E}$
非常滿意	157	0.36	$801 \times 0.36 = 288$	-131	$(-131)^2 = 17,161$	59.51
滿意	463	0.38	$801 \times 0.38 = 304$	159	$(151)^2 = 25,281$	83.06
尚可	118	0.19	$801 \times 0.19 = 152$	-34	$(-34)^2 = 1,156$	7.6
不滿意	20	0.03	$801 \times 0.03 = 24$	-4	$(-4)^2 = 16$	0.67
非常不滿意	43	0.04	$801 \times 0.04 = 32$	11	$(11)^2 = 121$	3.78
						154.61

配合度檢定

表15.6 旅客滿意程度的觀察次數與理論次數

	A	B	C	D	E	F	G	H	I	J	K
1	觀察次數										
2	非常滿意	滿意	尚可	不滿意	非常不滿意	非常滿意	滿意	尚可	不滿意	非常不滿意	
3	157	463	118	20	43	288	304	152	24	32	

配合度檢定

圖15.7 卡方檢定的對話方塊



獨立性檢定

○ 獨立性檢定意義

獨立性檢定是用來檢定兩個屬性間是否獨立的統計方法。

○ 獨立性檢定的檢定統計量

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

其中： r ：列聯表中橫列的個數， c ：縱行的個數， O_{ij} ：樣本觀察次數， \hat{E}_{ij} ：估計的理論次數(下面的例子會說明其計算方法)。自由度為 $(r-1)(c-1)$ 。

獨立性檢定

○ 決策法則

- ①若 $\chi^2 > \chi^2_{(c-1)(r-1), \alpha}$ ，則拒絕 H_0 。
- ②若 $\chi^2 \leq \chi^2_{(c-1)(r-1), \alpha}$ ，則接受 H_0 。

獨立性檢定

表15.7 台灣股價漲跌與外資買賣起

	上漲週數	下跌週數	總週數
買超週數	48	31	79
賣超週數	6	19	25
總週數	54	50	104

獨立性檢定

表15.8 台灣股價漲跌與外資買賣起

	上漲週數	下跌週數	總週數
買買超(大盤漲跌)	48 (41)	31 (38)	79
買超週數	6 (13)	19 (12)	25
總週數	54	50	104

齊一性檢定

○ 齊一性檢定意義

齊一性檢定是檢定兩個或兩個以上母體的某一特性的分配(各類別的比例)是否齊一或相近。

○ 齊一性檢定統計量

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

其中： r ：橫列個數， c ：縱行個數， O_{ij} ：樣本觀察次數， \hat{E}_{ij} ：估計期望論次數，自由度為 $(r-1)(c-1)$ 。

○ 決策法則

- ①若 $\chi^2 > \chi^2_{(c-1)(r-1), \alpha}$ ，則拒絕 H_0 。
- ②若 $\chi^2 \leq \chi^2_{(c-1)(r-1), \alpha}$ ，則接受 H_0 。

齊一性檢定

表15.9 廣告態度與年齡層

年齡	喜歡廣告	厭惡廣告	列總和
18歲以下	6	16	22
19~20歲	26	61	87
21~22歲	33	42	75
23~25歲	9	6	15
25歲以上	3	2	5
行總和	77	127	204

齊一性檢定

表15.10 外語測驗成績的人數分佈

分數	第一次	第二次	合計人數
80~100	10	7	17
60~79	21	23	44
40~59	11	9	20
0~39	4	4	8
合計人數	46	43	89

齊一性檢定

表15.11 廣告態度與年齡層的理论次數與觀察次數

	喜歡廣告	厭惡廣告	列總和
18歲以下	6 (8.30)	16 (13.70)	22
19~20歲	26 (23.84)	61 (54.16)	87
21~22歲	33 (28.31)	42 (46.69)	75
23歲以上	12 (7.55)	8 (12.45)	20
行總和	77	127	204

齊一性檢定

表15.12 外語能力考試的成績分配

分數	第一次	第二次	合計人數
80~100	10 (8.79)	7 (8.21)	17
60~79	21 (22.74)	23 (21.26)	44
40~59	11 (10.34)	9 (9.66)	20
0~39	4 (4.13)	4 (3.87)	8
合計人數	$n_{.} = 46$	$n_{.} = 43$	89

齊一性檢定

表15.13 大學聯考成績的分配

大學聯考結果	高中入學方式		小計
	聯考生	非聯考生	
錄取	31	24	55
未錄取	19	17	36
小計	50	41	91

齊一性檢定

表15.14 大學聯考成績的次數分配

大學聯考結果	高中入學方式		小計
	聯考生	非聯考生	
錄取	31 (30.2)	24 (24.8)	55
未錄取	19 (19.8)	17 (16.2)	36
小計	50	41	91

林惠玲 陳正倉著 雙葉書廊發行 2008

EXCEL 的使用

表15.15 選舉人名單

	A	B	C	D	E	F
1	編號	姓名	性別	年齡	教育程度	政黨傾向
2	1	黃文章	男	33	大學	K
3	2	李麗娟	女	56	高中	M
4	3	陳文玲	女	36	研究所	K
5	4	徐小翠	女	62	國中	M
6	5	胡光雄	男	57	高中	K
7	6	朱天文	男	26	大學	K
8	7	鍾典紅	女	59	國中	M

林惠玲 陳正倉著 雙葉書廊發行 2008

EXCEL 的使用

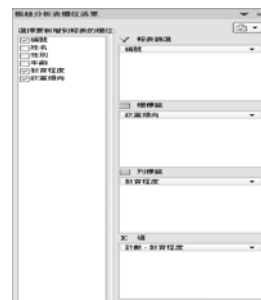
圖15.6 資料來源



林惠玲 陳正倉著 雙葉書廊發行 2008

EXCEL 的使用

圖15.7 分配資料於適當欄位



林惠玲 陳正倉著 雙葉書廊發行 2008

EXCEL 的使用

表15.16 教育程度與政黨取向

計數 - 教育欄標籤	教育欄標籤		總計
列標籤	K	M	
大學	2		2
研究所	1		1
高中	1	1	2
國中		2	2
總計	4	3	7

林惠玲 陳正倉著 雙葉書廊發行 2008