

統計函數(二)

標準常態分配反函數

- NORMSINV(累計機率)
- 於標準常態分配 ($\mu = 0$, $\sigma = 1$)，求某累計機率所對應之 z 值

B5		fx =NORMSINV(A5)			
	A	B	C	D	E
1	標準常態分配，均數為0，標準差為1				
2					
3	NORMSINV()				
4	自左尾累積	z			
5	0.001	-3.09			
6	0.025	-1.96			
7	0.050	-1.64			
8	0.100	-1.28			
9	0.250	-0.67			
10	0.500	0.00			
11	0.600	0.25			
12	0.750	0.67			
13	0.900	1.28			
14	0.950	1.64			
15	0.975	1.96			
16	0.990	2.33			
17					

信賴區間

- CONFIDENCE(α, σ, n)
- 可傳回母體平均數的信賴區間， α 為顯著水準， $\alpha=0.05$ 時表示求算95%之信賴區間。 σ 為母體標準差， n 為樣本數
- 若處理對象為常態分配，母體標準差（ σ ）已知，其計算公式：

$$\bar{x} \pm z\left(\frac{\alpha}{2}\right) \frac{\sigma}{\sqrt{n}}$$

信賴區間

- 學生通學之距離為常態分配，其標準差為3.2公里。今抽取100為同學為樣本，其通學之距離均數為8.6公里。由於 $z(0.25\%)$ 之值為1.96，故通學距離之95%信賴區間的公式：
- 其結果為7.9728~9.2272，表示有95%的學生之通學距離是介於此一範圍內
- 以CONFIDENCE()表示時，其下限及上限之公式應分別為：
=A2-CONFIDENCE(D2,B2,C2)
=A2+CONFIDENCE(D2,B2,C2)
- 【[CONF](#)】工作表

$$8.6 \pm 1.96 \frac{3.2}{\sqrt{100}}$$

C4		fx =A2+CONFIDENCE(D2,B2,C2)				
	A	B	C	D	E	F
1	樣本均數	標準差	樣本數	α		
2	8.6	3.2	100	5%		
3						
4	信賴區間	7.9728	9.2272			
5						

信賴區間

- 實務上，很少會已知母體標準差，就以樣本標準差來替代。
- 資料先以AVERAGE()、STDEV()與COUNT()求得均數、標準差與樣本數。
- 再以CONFIDENCE()求信賴區間
- 其下限及上限之公式應分別為：
 - =C8-CONFIDENCE(F8,D8,E8)
 - =C8+CONFIDENCE(F8,D8,E8)
- 【[CONF1](#)】工作表

E12		fx =C8+CONFIDENCE(F8,D8,E8)					
	A	B	C	D	E	F	G
7	分數		均數	標準差	樣本數	α	
8	85		66.33	22.53	15	5%	
9	25						
10	60						
11	80		均數的95%信賴區間				
12	87		54.93	到	77.74		
13	90						
14	80						
15	32						
16	80						
17	45						
18	85						
19	52						
20	88						
21	38						
22	68						
23							

Z檢定

- **ZTEST(數列,x,σ)**
- 傳回雙尾z檢定之P值（常態分配的雙尾機率值）
- 數列是要檢定相對於x之陣列或資料範圍
- x是要檢定之數值
- σ是母群體（已知）的標準差。若省略，則使用樣本標準差。

Z檢定

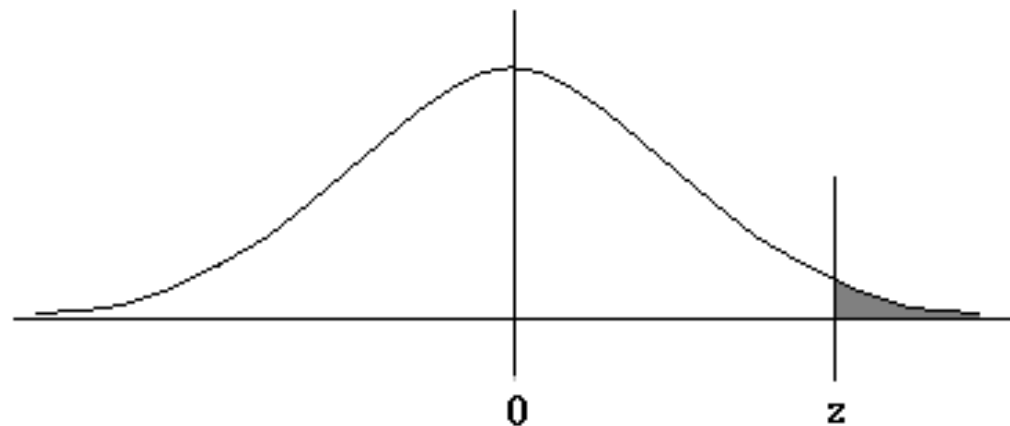
- 本函數之公式為：
$$Z = \frac{\mu - x}{\sigma / \sqrt{n}}$$

- $ZTEST(\text{陣列}, x, \sigma) = 1 - \text{NORMSDIST}(z)$

- z 即常態分配之 z 值

- 故本函數即算出：

1 - 自標準常態分配的左尾累加到 z 值處的機率
即圖之右尾的機率



Z檢定

- 如：ZTEST(數列,x)表要檢定母體均數（ μ ）是否等於抽樣之數列的均數（ x ）
- 其虛無假設與對立假設為：
 - $H_0 : \mu = x$
 - $H_1 : \mu \neq x$
- 判斷檢定結果時很簡單，只須看此P值是否小於您所指定之 α 值（顯著水準）的一半（雙尾檢定之故）。
- 如：P若為 $0.014 < (\alpha = 0.05/2)$ ，即表示在 $\alpha = 0.05$ 時，此檢定結果要棄卻虛無假設，接受對立假設 $\mu \neq x$ 。

Z檢定

- A1:F1為自全班隨機抽取幾位學生之成績（按理，應抽取30個以上之樣本，才適用z-test。
- 在此只是象徵性的抽幾個，問是可否接受全班成績為80分之假設？
- 檢定結果之P值為0.664624，大於0.025。即表示在 $\alpha=0.05$ 時，此檢定結果無法棄卻（fail to reject）虛無假設（ $\mu=x$ ），也就是無法棄卻全班成績之均數為80之假設
- 【ZTEST】工作表

B4		fx =ZTEST(A1:F1,B3)					
	A	B	C	D	E	F	G
1	75	78	79	80	82	83	
2							
3	X	80					
4	ZTEST	0.664624					
5							

Z檢定

- 若將X改為76，則此一檢定之結果為：
- 檢定結果之P值為0.001，小於0.025。
- 即表示在 $\alpha=0.05$ 時，此檢定結果要棄卻虛無假設（ $\mu=x$ ），也就是否定全班成績均數為76之假設。

B4		fx =ZTEST(A1:F1,B3)					
	A	B	C	D	E	F	G
1	75	78	79	80	82	83	
2							
3	X	76					
4	ZTEST	0.001					
5							

Z檢定

- 於全年級**300**位學生中，抽取下列幾個學生之成績，是否可證明全年級之母體均數恰為**70**分？（ $\alpha=0.05$ ）
- 答案：棄卻母體均數恰為**70**分之虛無假設
- 詳[ZTEST](#)工作表

	A	B	C	D	E
1	67	72	75	73	
2	49	75	88	75	
3	58	63	67	55	
4	36	77	75	72	
5	74	83	82	77	
6	89	65	86	80	
7	81	70	67	60	
8	86	72	84	91	
9	96	70	81	60	
10	96	75	82	78	
11	78	89	65	46	
12	85	83	38	72	
13					

Z檢定-兩個母體平均數差異檢定

- 甲班之母體變異數為4.5；乙班之母體變異數為12.5。
- 各隨機抽取9位同學之成績如表
- 試檢定兩班之平均成績是否存有顯著差異？
- 【[成績](#)】工作表

C12		fx		12.5
	A	B	C	D
1		甲班	乙班	
2		75	65	
3		78	67	
4		79	72	
5		80	81	
6		82	90	
7		83	55	
8		82	44	
9		90	62	
10		64	48	
11				
12	母體變異數	4.5	12.5	
13				

Z檢定-兩個母體平均數差異檢定

- 選「Z檢定-兩個母體平均數差異檢定」，續按〔確定〕鈕
- 於『變數1的範圍』與『變數2的範圍』設定兩組資料之範圍（B1:B10與C1:C10）
- 於『假設的均數差』輸入0，兩均數若相等其差為0
- 於『變數1之變異數』與『變數2之變異數』處，輸入已知之母體變異數（4.5與12.5）
- 點選「標記(L)」（因兩組資料均含『甲班』、『乙班』之字串標記）

z-檢定：兩個母體平均數差異檢定

輸入

變數 1 的範圍(1):

變數 2 的範圍(2):

假設的均數差(P):

變數 1 之變異數(已知)(V):

變數 2 之變異數(已知)(A):

☒ 標記(L)

α (A):

輸出選項

☒ 輸出範圍(O):

☐ 新工作表(P):

☐ 新活頁簿(W)

z檢定-兩個母體平均數差異檢定

- α 維持0.05
- 設定輸出範圍：目前工作表之E2位置
- 按〔確定〕鈕，結束
- 若對立假設為： $H_1: \mu_1 > \mu_2$
- 則應使用單尾檢定，依此結果：z值10.42903>單尾臨界值1.644853，故可知甲班之均數明顯高於乙班之均數。
- 由於本例僅在檢定其是否相等，故為一雙尾檢定。依此結果：z值10.42903>雙尾臨界值1.959961，故可知甲乙班之均數並不相等（存有顯著差異）。

z-檢定：兩個母體平均數差異檢定

輸入

變數 1 的範圍(1): \$B\$1:\$B\$10

變數 2 的範圍(2): \$C\$1:\$C\$10

假設的均數差(P): 0

變數 1 之變異數(已知)(V): 4.5

變數 2 之變異數(已知)(A): 12.5

☒ 標記(L)

α (A): 0.05

輸出選項

☒ 輸出範圍(O): \$E\$1

☐ 新工作表(P):

☐ 新活頁簿(W)

確定 取消 說明(H)

	A	B	C	D	E	F	G
1		甲班	乙班		z 檢定：兩個母體平均數差異檢定		
2		75	65				
3		78	67			甲班	乙班
4		79	72		平均數	79.22222	64.88889
5		80	81		已知的變異數	4.5	12.5
6		82	90		觀察值個數	9	9
7		83	55		假設的均數差	0	
8		82	44		z	10.42903	
9		90	62		P(Z<=z) 單尾	0	
10		64	48		臨界值：單尾	1.644854	
11					P(Z<=z) 雙尾	0	
12	母體變異數	4.5	12.5		臨界值：雙尾	1.959964	

Z檢定-兩個母體平均數差異檢定

- 像此種檢定，並不限定兩組之樣本數必須相等。
- 如檢定結果，就顯示因z值 $1.35 < \text{雙尾臨界值 } 1.96$ ，並無法棄卻甲乙班均數並無顯著差異之虛無假設（ $H_0: \mu_1 = \mu_2$ ）。
- **【[成績1](#)】** 工作表

E2		fx		z 檢定：兩個母體平均數差異檢定			
	A	B	C	D	E	F	G
1		甲班	乙班				
2		75	82		z 檢定：兩個母體平均數差異檢定		
3		78	67				
4		79	80			甲班	乙班
5		80	81		平均數	79.85714	78.33333
6		82	90		已知的變異數	4.5	5.6
7		83	81		觀察值個數	7	9
8		82	75		假設的均數差	0	
9			62		z	1.35	
10			87		P(Z<=z) 單尾	0.09	
11					臨界值：單尾	1.64	
12	母體變異數	4.5	5.6		P(Z<=z) 雙尾	0.18	
13					臨界值：雙尾	1.96	
14							

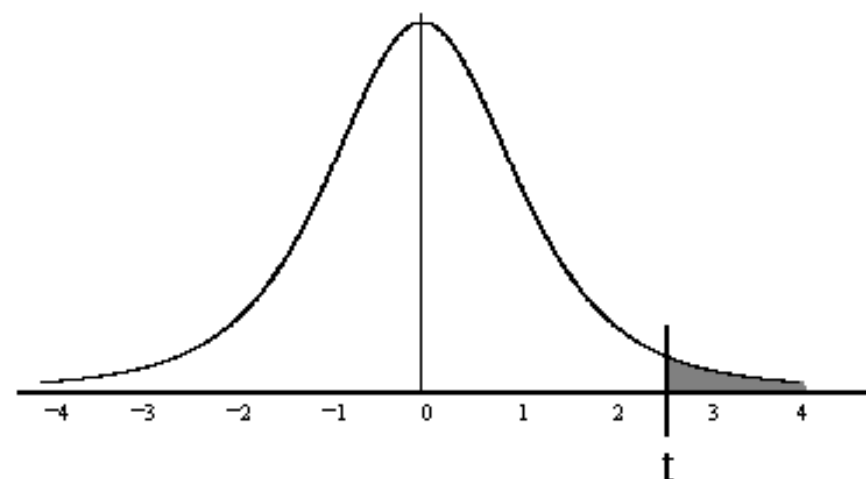
Z檢定-兩個母體平均數差異檢定

- 於兩地區分別隨機抽取**15**位樣本之所得。
- 分別以**VAR()**所求變異數為其母體變異數。
- 試於 $\alpha=0.05$ 之顯著水準下
 - 檢定兩地區之平均所得是否存有顯著差異？
 - 北區是否較南區高？
- 答案：*存有顯著差異，北區平均較南區高*
- **【所得】** 工作表

	A	B	C	D
1		北區	南區	
2		42,500	43,000	
3		37,000	45,000	
4		28,000	22,000	
5		52,000	32,000	
6		64,500	36,000	
7		18,750	35,000	
8		62,000	28,000	
9		37,000	27,000	
10		43,000	34,000	
11		58,500	35,000	
12		62,000	38,000	
13		74,000	39,000	
14		28,000	42,000	
15		52,600	46,000	
16		44,320	41,000	
17				

t分配

- **TDIST(t,自由度,單尾或雙尾)**
- **t**是要用來計算累計機率之**t**值。
- 自由度，**t**分配之自由度為樣本數減**1**，**n-1**
- 單尾或雙尾指定要傳回單尾或雙尾之累計機率值？
 - 為**1**，表傳回單尾之累計機率值
 - 為**2**，表傳回雙尾之累計機率值
- 本函數在求：於某一自由度下之**t**分配中，求**t**值以外之右尾的總面積（機率）。
 - 如為單尾，即求圖之陰影部份
 - 如為雙尾，即求左右兩尾之陰影部份



t分配

- t分配之圖形及機率值，將隨自由度不同而略有不同。
- 以自由度為**10**之情況下，不同t值所求得之單尾及雙尾累計機率分別為：
- **【TDIST】** 工作表

B2		fx =TDIST(A2,10,1)				
	A	B	C	D	E	F
1	t值	單尾	雙尾			
2	0.00	50.0%	100.0%	← =TDIST(A2,10,2)		
3	0.50	31.4%	62.8%	← =TDIST(A3,10,2)		
4	0.70	25.0%	50.0%	← =TDIST(A4,10,2)		
5	1.37	10.0%	20.1%	← =TDIST(A5,10,2)		
6	1.81	5.0%	10.0%	← =TDIST(A6,10,2)		
7	2.23	2.5%	5.0%	← =TDIST(A7,10,2)		
8	2.76	1.0%	2.0%	← =TDIST(A8,10,2)		
9	3.17	0.5%	1.0%	← =TDIST(A9,10,2)		
10						

t分配反函數

- **TINV(累計機率,自由度)**
- 於已知自由度之t分配中，求某累計機率所對應之t值。
 - 此t為依雙尾累計機率所求
 - 若要求單尾之t值，得將累計機率乘以2
- 由於t分配之圖形及機率值，將隨自由度不同而略有不同。
- 以自由度為10之情況下，所求得之結果。
- **【TINV】** 工作表

[illegible]

t分配反函數

- 練習：
 - 查d.f.為1~15之情況下，單尾機率為25%、10%、5%、2.5%、1%與0.5%之t值
- 【[TINV1](#)】工作表

	A	B	C	D	E	F	G	H
1		右尾機率						
2	df.	25%	10%	5%	2.5%	1%	0.5%	
3	1	1.00	3.08	6.31	12.71	31.82	63.66	
4	2	0.82	1.89	2.92	4.30	6.96	9.92	
5	3	0.76	1.64	2.35	3.18	4.54	5.84	
6	4	0.74	1.53	2.13	2.78	3.75	4.60	
7	5	0.73	1.48	2.02	2.57	3.36	4.03	
8	6	0.72	1.44	1.94	2.45	3.14	3.71	
9	7	0.71	1.41	1.89	2.36	3.00	3.50	
10	8	0.71	1.40	1.86	2.31	2.90	3.36	
11	9	0.70	1.38	1.83	2.26	2.82	3.25	
12	10	0.70	1.37	1.81	2.23	2.76	3.17	
13	11	0.70	1.36	1.80	2.20	2.72	3.11	
14	12	0.70	1.36	1.78	2.18	2.68	3.05	
15	13	0.69	1.35	1.77	2.16	2.65	3.01	
16	14	0.69	1.35	1.76	2.14	2.62	2.98	
17	15	0.69	1.34	1.75	2.13	2.60	2.95	
18								

t檢定

- **TTEST(第一組資料,第一組資料,單尾或雙尾,類型)**
- t檢定是用來進行兩組小樣本（ $n < 30$ ）資料之均數檢定或成對樣本的均數差檢定。
- 除成對樣本外，兩組資料之樣本數允許不同。
- 單尾或雙尾是以1或2來標示，其類型則可分為下列三種：
 - 1 成對
 - 2 具有相同變異數的二個樣本
 - 3 具有不同變異數的二個樣本

t檢定

- 通常，於實務上，要事先以F檢定，判定兩群體之變異數是否相同。才可決定應使用何種類型之t檢定。
 - （參見本章『F檢定FTEST()』處之說明）
- 本函數所回應之值為其右尾之機率（P），判斷檢定結果時很簡單
 - 只須看此P值是否小於所指定顯著水準之 α 值（單尾）；
 - 或 α 值的一半（雙尾）

兩獨立樣本均數檢定(變異數相同)

- 兩樣本平均數的t檢定
 - 旨在比較變異數相同的兩個母群之間平均數的差異
 - 或比較來自同一母群之兩個樣本之間的不同
- 若兩母群體之變異數相同（ $\sigma_1^2 = \sigma_2^2$ ），是採用匯總變異數t檢定(pooled-variance t test)
- 其相關公式為：

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_p^2}{N_1} + \frac{S_p^2}{N_2}}}$$

$$S_p^2 = \frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}$$

$$d.f. = N_1 + N_2 - 2$$

式中， S_p^2 即匯總變異數

兩獨立樣本均數檢定(變異數相同)

- 班上男女生之成績的變異數相同。
- 依抽樣資料，是否可證明，在 $\alpha=0.05$ 之顯著水準下，男女生之成績無差異存在？
- 由於是變異數相同，t檢定之類型為2。
且虛無假設與對立假設分別為：
 - $H_0: \mu_1 = \mu_2$
 - $H_1: \mu_1 \neq \mu_2$
 - 故此類檢定為雙尾
- B15公式=TTEST(B2:B11,C2:C9,2,2)
- 【TTEST】工作表

B15		fx =TTEST(B2:B11,C2:C9,2,2)				
	A	B	C	D	E	F
1		男	女			
2		76	81			
3		74	82			
4		70	78			
5		80	85			
6		68	79			
7		90	81			
8		72	82			
9		75	85			
10		78				
11		72				
12						
13	平均	75.5	81.6			
14						
15	T-test	0.019306	<-- =TTEST(B2:B11,C2:C9,2,2)			
16						

兩獨立樣本均數檢定(變異數相同)

- 由其TTEST()函數所獲得之P值為 $0.019 < \alpha = 0.05$ ，故將棄卻男女成績相等之虛無假設。
- 也就是說，在 $\alpha = 0.05$ 之顯著水準下，男女生之成績存有顯著差異。

B15		fx		=TTEST(B2:B11,C2:C9,2,2)		
A	B	C	D	E	F	
1	男	女				
2	76	81				
3	74	82				
4	70	78				
5	80	85				
6	68	79				
7	90	81				
8	72	82				
9	75	85				
10	78					
11	72					
12						
13	平均	75.5	81.6			
14						
15	T-test	0.019306	<-- =TTEST(B2:B11,C2:C9,2,2)			
16						

兩獨立樣本均數檢定(變異數相同)

- 點選「標記(L)」(因兩組資料均含『男』、『女』之字串標記)
- α 維持0.05
- 設定輸出範圍，本例安排於目前工作表之E2位置
- 按〔確定〕鈕，結束
- 由於本例僅在檢定其是否相等，故為一雙尾檢定。
- 依此結果：自由度為16，t統計值之絕對值2.60088>雙尾臨界值2.119905 (F14處之P值0.019306< α =0.05，同於B15)
- 故得棄卻男女成績均數相等之虛無假設(兩者存有顯著差異)

	B	C	D	E	F	G	H
1	男	女					
2	76	81		t 檢定：兩個母體平均數差的檢定，假設變異數相等			
3	74	82					
4	70	78			男	女	
5	80	85		平均數	75.5	81.625	
6	68	79		變異數	38.94444	6.267857	
7	90	81		觀察值個數	10	8	
8	72	82		Pooled 變異數	24.64844		
9	75	85		假設的均數差	0		
10	78			自由度	16		
11	72			t 統計	-2.60088		
12				P(T<=t) 單尾	0.009653		
13	75.5	81.6		臨界值：單尾	1.745884		
14				P(T<=t) 雙尾	0.019306		
15	0.01930631	← =T.TEST(B2:B11,C2:C9,2,2)		臨界值：雙尾	2.119905		
16	0.01930631	← =TTEST(B2:B11,C2:C9,2,2)					
17							

t 檢定：兩個母體平均數差的檢定，假設變異數相等

輸入

變數 1 的範圍(1):

變數 2 的範圍(2):

假設的均數差(P):

☒ 標記(L)

α (A):

輸出選項

☒ 輸出範圍(O):

☐ 新工作表(P):

☐ 新活頁簿(W):

確定 取消 說明(H)

兩獨立樣本均數檢定(變異數相同)

- 練習：
 - 兩年度之所得變異數相同
 - 若 $\alpha=0.05$ ，資料是否表示1999年之每月所得明顯高過1997年：
(此為單尾檢定)
- 答案：無法棄卻兩年度之每月所得相等之虛無假設
- 【[所得](#)】工作表

	A	B	C	D
1	假定，兩年度每月所得變異數相同。			
2				
3		1997	1999	
4		45,038	35,700	
5		36,125	40,250	
6		38,751	28,600	
7		26,760	18,900	
8		18,500	60,200	
9		42,000	48,560	
10		38,000	37,800	
11		40,500	52,500	
12		28,790	53,400	
13		35,600	29,600	
14			32,105	
15			30,645	
16				

兩獨立樣本均數檢定（變異數不等）

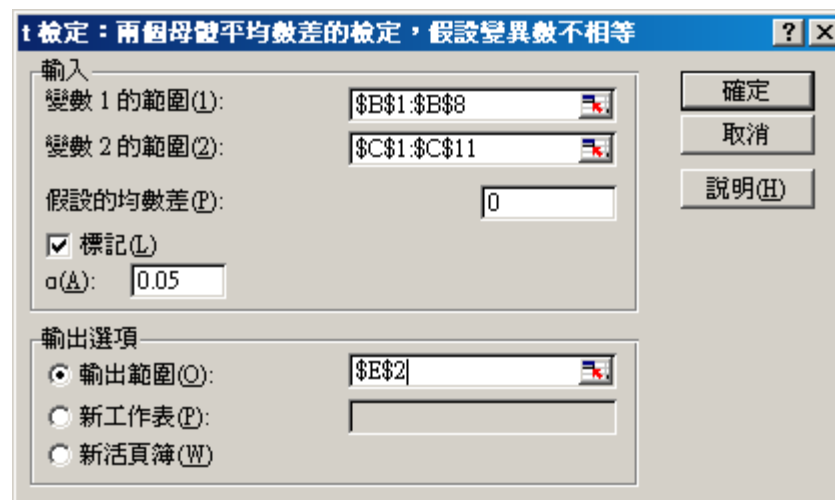
- 若兩母群體之變異數不同（ $\sigma_1^2 \neq \sigma_2^2$ ），則將用個別變異數的t統計量(Cochran & Cox法)。
- 請注意，其自由度已不再是兩母群體之變異數相等時簡單的d.f.=N1+N2-2，依此處公式計算之自由度可能會含小數。

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$$

$$d.f. = \frac{\left(\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2} \right)}{\frac{\left(\frac{S_1^2}{N_1} \right)^2}{(N_1 - 1)} + \frac{\left(\frac{S_2^2}{N_2} \right)^2}{(N_2 - 1)}}$$

兩獨立樣本均數檢定（變異數不等）

- 點選「標記(L)」（因兩組資料均含『甲班』、『乙班』之字串標記）
- α 維持0.05
- 設定輸出範圍：目前工作表E2位置
- 按〔確定〕鈕



The image shows the 't-Test: Two-Sample Assuming Unequal Variances' dialog box in Excel. The 'Input' section has 'Variable 1 Range' set to '\$B\$1:\$B\$8', 'Variable 2 Range' set to '\$C\$1:\$C\$11', and 'Hypothesized Mean Difference' set to '0'. The 'Marked in input range' checkbox is checked, and the significance level ' α ' is set to '0.05'. The 'Output Options' section has 'Output Range' set to '\$E\$2', with 'New Worksheet' and 'New Workbook' options unselected. Buttons for 'OK', 'Cancel', and 'Help' are on the right.

E2				t 檢定：兩個母體平均數差的檢定，假設變異數不相等				
A	B	C	D	E	F	G	H	I
1	甲班	乙班						
2		81	72	t 檢定：兩個母體平均數差的檢定，假設變異數不相等				
3		82	80					
4		92	68		甲班	乙班		
5		85	90	平均數	83.57143	77.4		
6		88	72	變異數	29.61905	37.6		
7		75	75	觀察值個數	7	10		
8		82	78	假設的均數差	0			
9			81	自由度	14			
10			78	t 統計	2.183118			
11			80	P(T<=t) 單尾	0.023274			
12				臨界值：單尾	1.761309			
13	平均	83.57	77.40	P(T<=t) 雙尾	0.046549			
14	假定兩班之變異數不同			臨界值：雙尾	2.144789			
15	T-Test	0.023						
16								

兩獨立樣本均數檢定（變異數不等）

- 由於本例是在檢定甲班平均是否大於乙班平均，故為一單尾檢定。
- 依此結果：自由度為**14**，**t**統計值**2.183**>單尾臨界值**1.761**（F11處之P值**0.023**< $\alpha=0.05$ ，同於B15）
- 故得棄卻甲乙班平均成績相等之虛無假設，接受甲班平均成績高於乙班平均成績之對立假設。

[illegible]

兩獨立樣本均數檢定（變異數不等）

- 練習：
 - 兩地區之所得變異數不同
 - 若 $\alpha=0.05$
 - 下列資料是否表示乙地區之所得明顯高過甲地區：（此為單尾檢定）
- 答案：乙地區之所得明顯高過甲地區
- 【[所得](#)】工作表

	A	B	C	D
1	假定，兩地區之所得變異數不同。			
2				
3		甲地	乙地	
4		35,700	48,760	
5		40,650	45,250	
6		28,600	60,200	
7		24,900	48,560	
8		40,200	37,800	
9		48,560	52,500	
10		27,800	73,400	
11		52,500	59,600	
12		36,700	32,105	
13			40,645	
14				

成對樣本

- 同組人，受訓後的打字速度是否高於受訓前
- 同一部車，左右使用不同廠牌輪胎，經過一段時間後，檢查其磨損程度，看甲廠牌之輪胎是否優於乙廠牌？...
- 諸如此類之例子，兩組受測樣本間為相依（同一個人、同一部車），就要使用配對樣本的t檢定。
- d即同一配對之兩資料相減之差

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$$

$$d = x_1 - x_2$$

$$d.f. = n - 1$$

成對樣本

- 抽7部車，左右使用不同廠牌輪胎，每車各由同一個人駕駛（同一駕駛習慣），經過一段時間後，獲得下示輪胎磨損之配對資料（千分之一吋）。
- 是否可證明，在 $\alpha=0.05$ 之顯著水準下，甲廠牌之輪胎磨損程度是否較乙廠牌大？
- 由於是配對樣本，t檢定之類型為1。且虛無假設與對立假設分別為：
 - $H_0: \mu_d = 0$
 - $H_1: \mu_d > 0$
 - 故此類檢定為單尾
 - B12公式=TTEST(B2:B8,C2:C8,1,1)
- 【[輪胎](#)】工作表

	B12	fx =TTEST(B2:B8,C2:C8,1,1)					
	A	B	C	D	E	F	G
1		甲廠	乙廠	d=甲-乙			
2		143	125	18			
3		68	64	4			
4		100	94	6			
5		35	38	-3			
6		105	90	15			
7		123	125	-2			
8		98	76	22			
9							
10	平均	96.00	87.43	8.57			
11							
12	T-Test	0.03	<-- =TTEST(B2:B8,C2:C8,1,1)				
13							

成對樣本

- 由其TTEST()函數所獲得之P值為 $0.03 < \alpha = 0.05$
- 故將棄卻甲乙廠之輪胎的耐磨程度相等之虛無假設
- 也就是說，在 $\alpha = 0.05$ 之顯著水準下，甲廠的輪胎磨損程度較乙廠大

B12		fx =TTEST(B2:B8,C2:C8,1,1)					
	A	B	C	D	E	F	G
1		甲廠	乙廠	d=甲-乙			
2		143	125	18			
3		68	64	4			
4		100	94	6			
5		35	38	-3			
6		105	90	15			
7		123	125	-2			
8		98	76	22			
9							
10	平均	96.00	87.43	8.57			
11							
12	T-Test	0.03	<-- =TTEST(B2:B8,C2:C8,1,1)				
13							

成對樣本

- 此例若要以TTEST()進行檢定，應另將訓練後之值-40置於另一新欄，再將其與訓練前之值進行配對檢定：
- 兩種方式，均顯示此套訓練並無法讓學生每分鐘多打40個字。
- 【打字1】工作表

B15		fx =TTEST(D2:D11,B2:B11,1,1)					
	A	B	C	D	E	F	G
1		訓練前	訓練後	訓練後-40			
2		12	53	13			
3		25	67	27			
4		18	60	20			
5		14	48	8			
6		23	72	32			
7		20	80	40			
8		24	65	25			
9		13	50	10			
10		12	47	7			
11		18	51	11			
12							
13	平均	17.9	59.3	19.3			
14							
15	T-Test	0.30	← =TTEST(D2:D11,B2:B11,1,1)				
16							

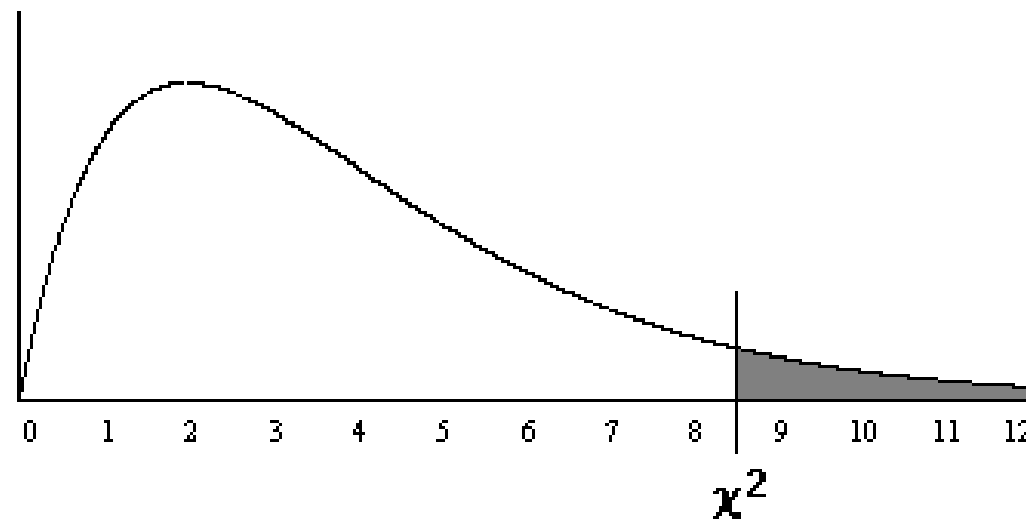
成對樣本

- 練習：
 - 要比較一套新減肥法之效果
 - 隨機抽取**12**位受測者進行測試一個月。
- 減肥前及減肥後之體重（公斤），於 $\alpha=0.05$ 之水準下，是否表示此套新減肥法可讓受測者至少減5公斤：（此為單尾檢定）
- 於 $\alpha=0.05$ 之水準下，此套新減肥法平均可讓受測者至少減5公斤
- 【減肥】工作表

	A	B	C	D
1		減肥前	減肥後	
2		56	50	
3		48	46	
4		62	56	
5		64	48	
6		51	47	
7		70	55	
8		53	48	
9		56	45	
10		49	47	
11		42	41	
12		70	68	
13		53	52	
14				

卡方分配

- CHIDIST(x,自由度)
- x是用來計算累計機率之卡方值 (χ^2)
- 自由度 (d.f.)，適合度檢定為組數減1 ($k-1$)、獨立性檢定或同質性檢定均為(列數-1) $\times$ (行數-1)即 $(r-1) \times (c-1)$ 。
- 本函數在求：
 - 於某一自由度下之卡方分配中，求x值以外之右尾的總面積（機率）。
 - 即傳回卡方分配之右尾累計機率值（陰影部份）



卡方分配

- 卡方分配之圖形及機率值，將隨自由度不同而略有不同。
- 以自由度為**10**與**20**之情況下，不同卡方值所求得之單尾累計機率分別為
- 【[卡方](#)】工作表

B3		fx		=CHIDIST(A3,\$B\$1)		
	A	B	C	D	E	F
1	d.f.	10		d.f.	20	
2	卡方值	右尾機率		卡方值	右尾機率	
3	12.55	0.250		23.82	0.250	
4	16.00	0.100		28.40	0.100	
5	18.30	0.050		31.40	0.050	
6	20.50	0.025		34.20	0.025	
7	23.20	0.010		37.60	0.010	
8	25.20	0.005		40.00	0.005	
9						

卡方分配反函數

- **CHIINV(累計機率,自由度)**
- 於已知自由度之卡方分配中，求某累計機率所對應之卡方值。
- 此卡方為依單尾累計機率所求。有了此函數。
- 由於卡方分配之圖形及機率值，將隨自由度不同而略有不同。下表是以自由度為**10**之情況下，所求得之結果：
- **【CHIINV】** 工作表

B4		fx		=CHIDIST(A4,\$B\$2)				
	A	B	C	D	E	F	G	H
1	CHIDIST()			CHIINV()				
2	d.f.	10		d.f.	10			
3	卡方值	右尾機率		右尾機率	卡方值			
4	12.55	0.250		25.0%	12.55	← =CHIINV(D4,\$E\$2)		
5	16.00	0.100		10.0%	16.00	← =CHIINV(D5,\$E\$3)		
6	18.30	0.050		5.0%	18.30	← =CHIINV(D6,\$E\$4)		
7	20.50	0.025		2.5%	20.50	← =CHIINV(D7,\$E\$5)		
8	23.20	0.010		1.0%	23.20	← =CHIINV(D8,\$E\$6)		
9	25.20	0.005		0.5%	25.20	← =CHIINV(D9,\$E\$7)		
10								

卡方分配反函數

- 練習：
- 查自由度（d.f.）為1~15之情況下，單尾機率为25%、10%、5%、2.5%、1%與0.5%之卡方值：
- 【[卡方值](#)】工作表

	A	B	C	D	E	F	G
1		右尾機率					
2	d.f.	25%	10%	5%	2.5%	1%	0.5%
3	1	1.32	2.71	3.84	5.02	6.63	7.88
4	2	2.77	4.61	5.99	7.38	9.21	10.60
5	3	4.11	6.25	7.81	9.35	11.34	12.84
6	4	5.39	7.78	9.49	11.14	13.28	14.86
7	5	6.63	9.24	11.07	12.83	15.09	16.75
8	6	7.84	10.64	12.59	14.45	16.81	18.55
9	7	9.04	12.02	14.07	16.01	18.48	20.28
10	8	10.22	13.36	15.51	17.53	20.09	21.95
11	9	11.39	14.68	16.92	19.02	21.67	23.59
12	10	12.55	15.99	18.31	20.48	23.21	25.19
13	11	13.70	17.28	19.68	21.92	24.73	26.76
14	12	14.85	18.55	21.03	23.34	26.22	28.30
15	13	15.98	19.81	22.36	24.74	27.69	29.82
16	14	17.12	21.06	23.68	26.12	29.14	31.32
17	15	18.25	22.31	25.00	27.49	30.58	32.80

卡方分配反函數

- 卡方之運算公式為：
- 即讓觀察值範圍的每一格減去期望值範圍對應位置之每一格的值，求平方，再除以期望值範圍對應位置之每一格的值，將這些值逐一加總即為卡方值：
- 其中，**r**為列數、**c**為欄數。然後計算自由度 **((r-1)*(c-1))**，並查表算出此一卡方值於卡方分配之右尾機率 (**P**)。

$$\chi^2 = \sum_{allcell} \frac{(O - E)^2}{E}$$

$$\chi^2 = \frac{(O_{1,1} - E_{1,1})^2}{E_{1,1}} + \frac{(O_{1,2} - E_{1,2})^2}{E_{1,2}} + \dots + \frac{(O_{r,c} - E_{r,c})^2}{E_{r,c}}$$

傳統計算卡方之方式

- 假定，對政府之某一項政策之民調結果，想以卡方檢定，判斷贊成與反對之比例高低，會不會因地區別之不同而有顯著差異。也就是要進行其關聯性檢定。
- 首先，求其欄/列加總與各合計佔全部人數之百分比，以方便於I4:K5計算各格之期望值。
 - 第一列第一欄之期望值應為：第一列% × 第一欄% × 總樣本數
 - 公式：=B\$7*\$F4*\$E\$6
- 【[民調](#)】工作表

傳統計算卡方之方式

	A	B	C	D	E
1	對某一政策之民調結果				
2	實際資料(觀察值範圍)				
3		北部	中部	南部	
4	贊成	200	75	40	
5	反對	50	70	140	
6					

	A	B	C	D	E	F	G	H	I	J	K	L
1	對某一政策之民調結果											
2	實際資料(觀察值範圍)											
3		北部	中部	南部	合計	%		期望值範圍				
4	贊成	200	75	40	315	54.8%		贊成	137.0	79.4	98.6	
5	反對	50	70	140	260	45.2%		反對	113.0	65.6	81.4	
6	合計	250	145	180	575	100.0%						
7	%	43.5%	25.2%	31.3%	100.0%							
8												

	A	B	C	D	E	F	G
1	對某一政策之民調結果						
2	實際資料(觀察值範圍)						
3		北部	中部	南部	合計	%	
4	贊成	200	75	40	315	54.8%	
5	反對	50	70	140	260	45.2%	
6	合計	250	145	180	575	100.0%	
7	%	43.5%	25.2%	31.3%	100.0%		
8							

傳統計算卡方之方式

- 將其安排於B11:D12
- B11之公式： $= (B4 - I4)^2 / I4$

$$\frac{(O - E)^2}{E}$$

$$\frac{(O_{1,1} - E_{1,1})^2}{E_{1,1}}$$

[illegible]

傳統計算卡方之方式

- 茲將各地區贊成與反對之人數，分別除以該地區之樣本數，求得下示之交叉表。
- 以B18與B20為例，其公式分別為：
 - =B4/B\$6
 - =B5/B\$6
- 由表中之資料可看出：
 - 整體上，贊成此一政策者仍居多數（54.8%）；
 - 贊成與反對之比例高低，會因地區別之不同而有顯著差異
 - 北部贊成者較多（80.0%）
 - 中部贊成與反對者差不多
 - 南部則是反對者較多（77.8%）

	B19		fx =B4/B\$6					
	A	B	C	D	E	F	G	H
17	某政策之民調結果							
18		北部	中部	南部	全體			
19	贊成	80.0%	51.7%	22.2%	54.8%			
20	反對	20.0%	48.3%	77.8%	45.2%			
21	合計	100.0%	100.0%	100.0%	100.0%			
22	樣本數	250	145	180	575			
23								
24	整體言，贊成者佔多數54.8%							
25	贊成與反對之比例高低，會因地區別之不同而有顯著差異							
26	北部贊成者較多（80.0%）							
27	中部贊成與反對者差不多							
28	南部則反對者較多（77.8%）							
29								

使用卡方檢定應注意事項

- 卡方檢定僅適用於類別資料（名目變數，如：性別、地區）
- 各細格之期望次數不應少於5
- 通常要有80%以上的期望次數 ≥ 5 ，否則會影響其卡方檢定的效果。
- 若有期望次數小於5時，可將其合併。

利用CHITEST()

- 求得觀察值範圍（B4:C6）與其期望值範圍（H4:I6）後
- 直接以
 - =CHITEST(B4:D5,I4:K5)
 - 利用CHITEST()函數求得其檢定結果（P值）
- 【CHITEST】工作表

[illegible]

利用CHITEST()

- 練習：
 - 以卡方檢定判斷各年齡層與經常求診之疾病項目間，是否毫無任何關聯性：
- 【[求診](#)】工作表

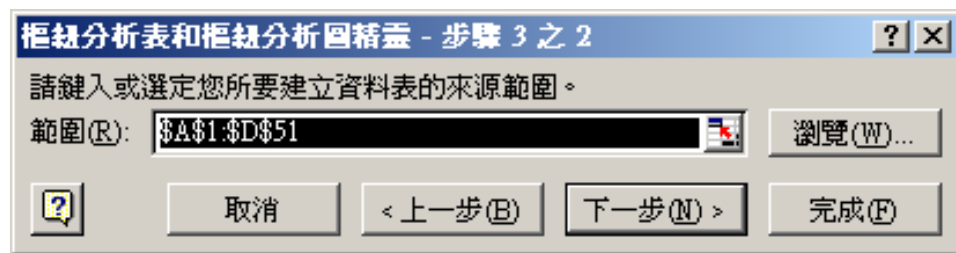
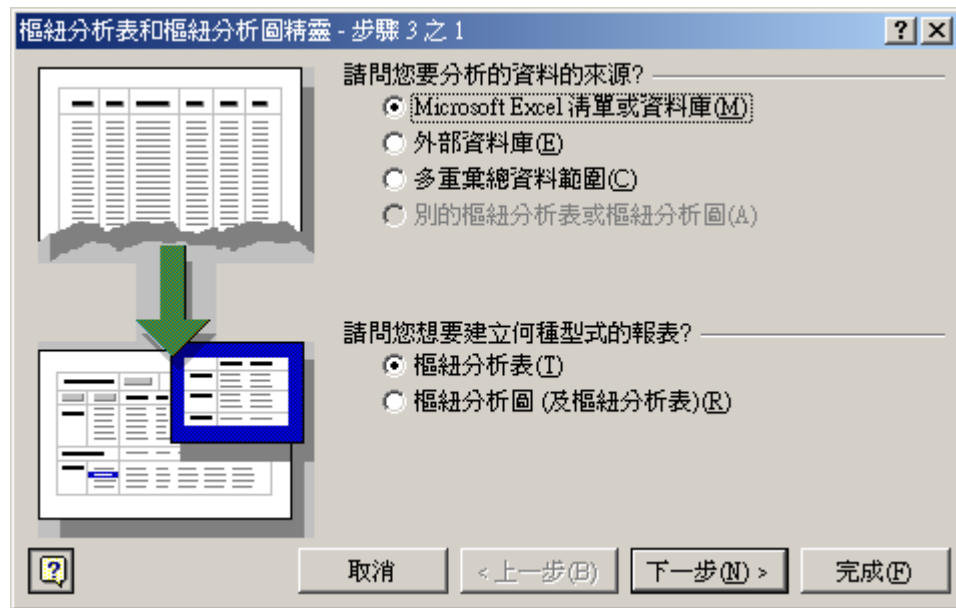
	A	B	C	D	E	F
1	年齡與疾病之關聯表					
2	實際資料(觀察)					
3		-30	30-45	45-60	60-	
4	眼	440	568	360	260	
5	鼻	960	1450	820	102	
6	喉	264	305	162	125	
7						

	A	B	C	D	E	F
9	期望值					
10		-30	30-45	45-60	60-	
11	眼	415.6	496.2	345.0	495.9	
12	鼻	203.1	242.4	168.6	242.3	
13	喉	790.5	943.6	656.1	943.1	
14						
15	直接以CHITEST求算之P值			0.00		
16	可知年齡不同，會影響疾病之類別					
17						

	A	B	C	D	E	F	G
18	年齡與疾病之關聯表						
19		-30	30-45	45-60	60-	合計	
20	眼	26.4%	24.5%	26.8%	53.4%	28.0%	
21	鼻	57.7%	62.4%	61.1%	20.9%	57.3%	
22	喉	15.9%	13.1%	12.1%	25.7%	14.7%	
23	合計	100.0%	100.0%	100.0%	100.0%	100.0%	
24	樣本數	1664	2323	1342	487	5816	
25							
26	整體言，半數以上之(57.3%)受訪者均患有鼻科疾病者						
27	60歲以上者患眼科及喉科之疾病者較多						
28	60歲以下者患鼻科之疾病者較多						
29							

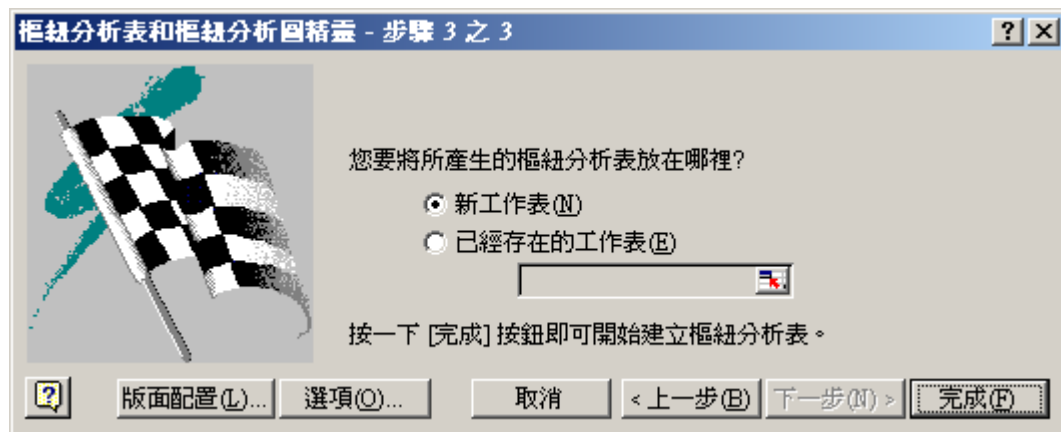
交叉分析表

- 以滑鼠單按問卷資料之任一儲存格
- 執行「**資料(D)/樞紐分析表及圖報表(P)**」，轉入『**樞紐分析表和樞紐分析圖精靈-步驟3之1**』對話方塊
- 於上半部之來源，選「**Microsoft Excel清單或資料庫(M)**」；於下半部之報表類型，選「**樞紐分析表(T)**」
- 按〔**下一步**〕鈕，轉入『**樞紐分析表和樞紐分析圖精靈-步驟3之2**』對話方塊，以標定建表之來源範圍



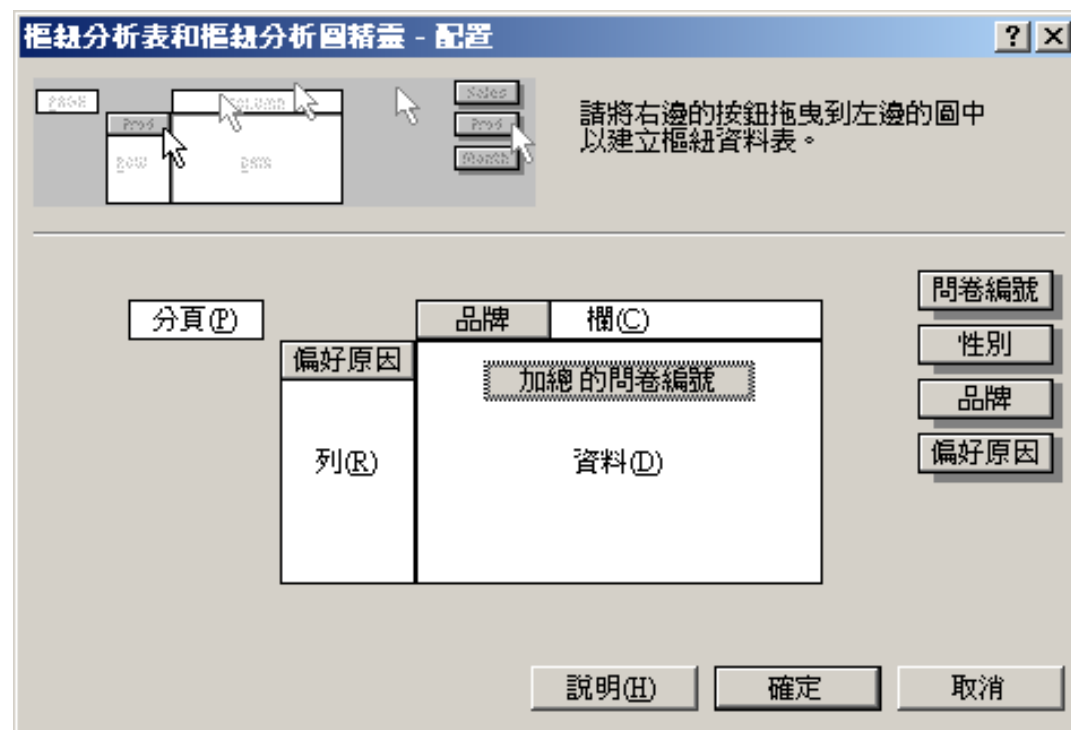
交叉分析表

- 按〔下一步〕鈕，轉入『樞紐分析表和樞紐分析圖精靈-步驟3之3』對話方塊
- 按〔下一步〕鈕，轉入『樞紐分析表和樞紐分析圖精靈-配置』對話方塊，以安排樞紐分析表之列、欄與資料內容



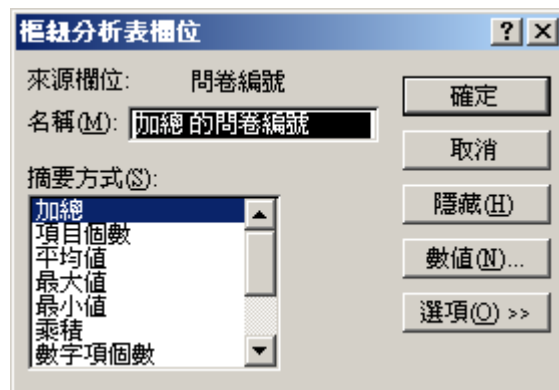
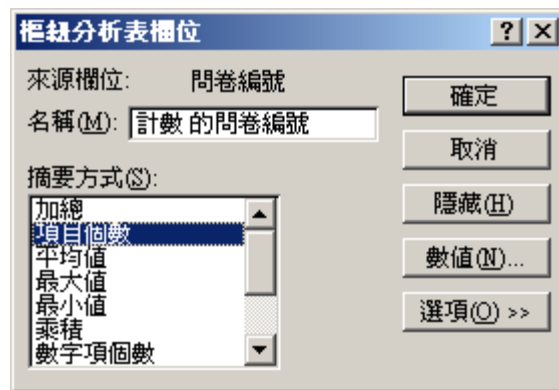
交叉分析表

- 欲先分析消費者使用某一『品牌』之『偏好原因』的交叉表
- 將拖曳到『欄(C)』處，拖曳到『列(R)』處，拖曳到『資料(D)』處
- 因為，係要求人數（即項目個數），故將那一個欄位拖曳到『資料(D)』處，均不影響本例之求算動作。



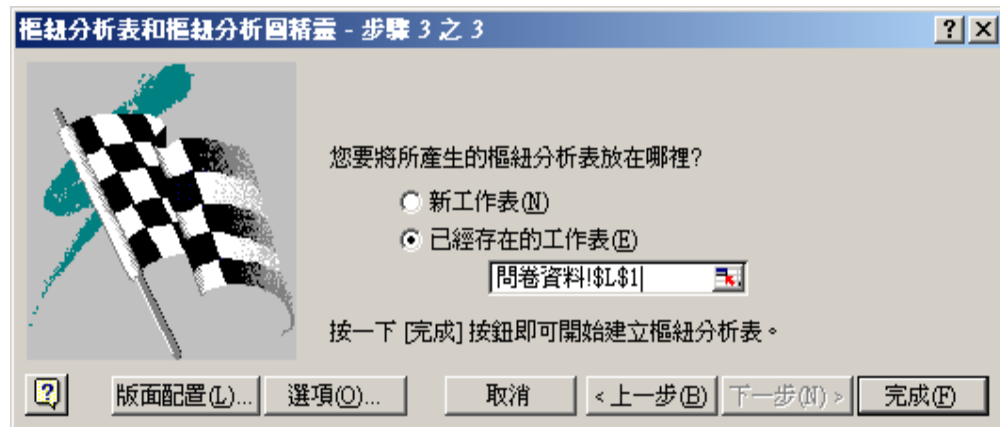
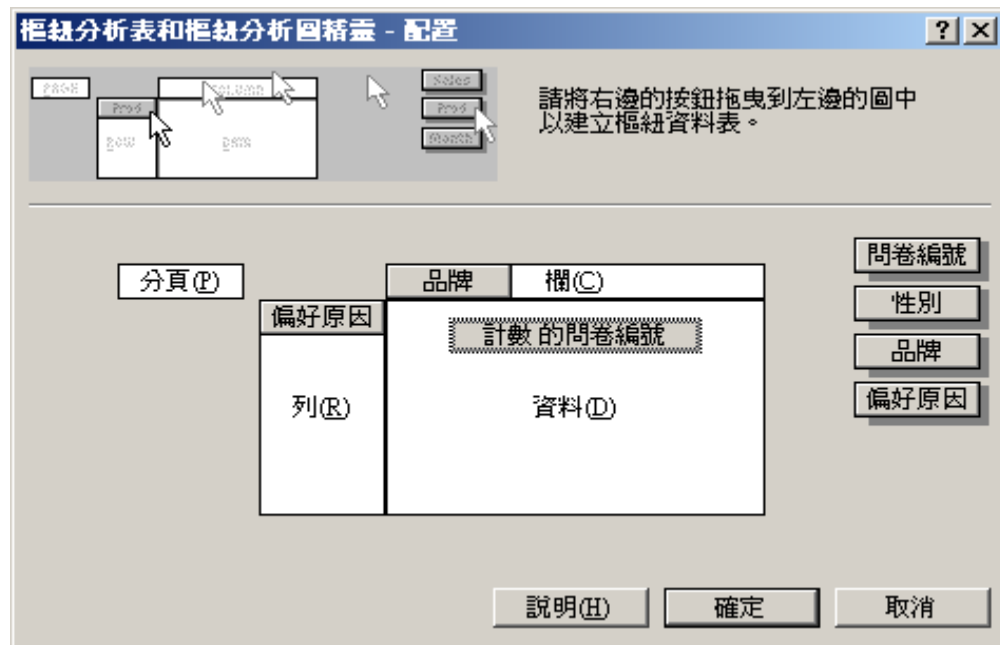
交叉分析表

- 由於表內『資料(D)』處所求算之統計量，預設為求算選取欄位之**加總**。並不符合我們的要求，因為是要求算其出現次數（即記錄數）。
- 故於其上雙按滑鼠
- 於『摘要方式(S):』處將其改為「項目個數」，以求算出現次數



交叉分析表

- 按〔確定〕鈕，『資料(D)』處所求算之統計量已改為計數
- 按〔確定〕鈕，回『樞紐分析表和樞紐分析圖精靈-步驟3之3』對話方塊，以安排要將樞紐分析表放在何處？
- 欲將樞紐分析表安排於目前工作表之L1處
- 先選「已經存在的工作表(E)」項，續選按L1儲存格



交叉分析表

- 按〔完成〕鈕，於L1處獲致交叉分析表
- 於M2:O2輸入該編號所對應之品牌名稱
- 於L3:L5輸入該編號所對應之偏好原因
- 於L1輸入『品牌交叉偏好原因』文字
- 以利閱讀

L1	L	M	N	O	P	Q	R	S
1	計數的問卷編號	品牌						
2	偏好原因	1	2	3	總計			
3		1	11	1	6	18		
4		2	2	8	5	15		
5		3	4	2	11	17		
6	總計	17	11	22	50			
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								

樞紐分析表

樞紐分析表(P) ▾

樞紐分析表欄位清單

拖曳項目到樞紐分析表

- ☐ 問卷編號
- ☐ 性別
- ☐ 品牌
- ☐ 偏好原因

新增至 列區域 ▾

L1	L	M	N	O	P	Q
1	品牌交叉偏好原因	品牌				
2	偏好原因	A牌	B牌	C牌	總計	
3	1. 價格便宜	11	1	6	18	
4	2. 品質優良	2	8	5	15	
5	3. 外型美觀	4	2	11	17	
6	總計	17	11	22	50	
7						

交叉分析表

- 可以計算期望值交叉表
- 並續而求算卡方值及其顯著水準。
- 由M20之P值為 $0.001 < \alpha = 0.05$ ，可知各品牌與其使用者偏好原因間，存有顯著之關聯性。

M20		fx =CHITEST(M3:O5,M11:O13)				
	L	M	N	O	P	Q
9	期望值	品牌				
10	偏好原因	A牌	B牌	C牌		
11	1. 價格便宜	19.8	0.3	4.5		
12	2. 品質優良	0.8	19.4	3.8		
13	3. 外型美觀	2.8	1.1	16.2		
14						
15	求算卡方	3.8912	2.2125	0.4655		
16		1.8843	6.6939	0.3879		
17		0.5482	0.8095	1.6565		
18						
19	卡方	18.55				
20	P值	0.001				
21						

交叉分析表

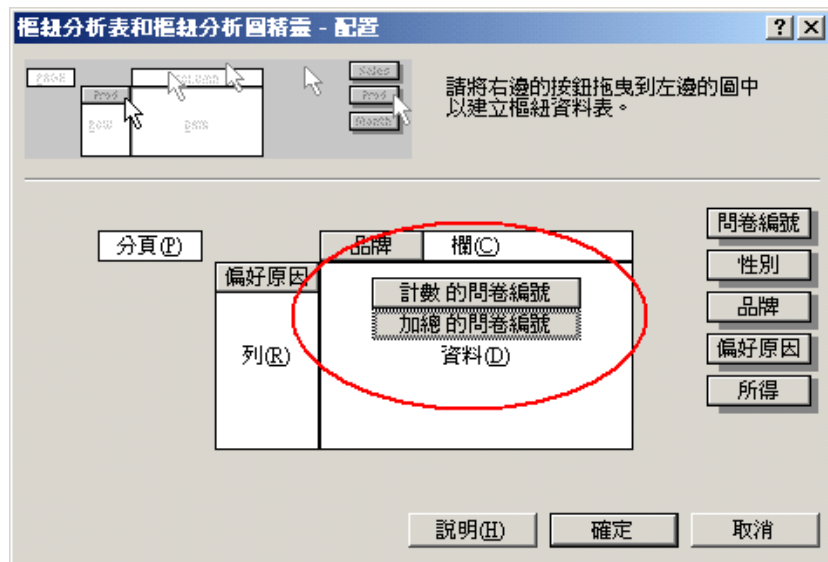
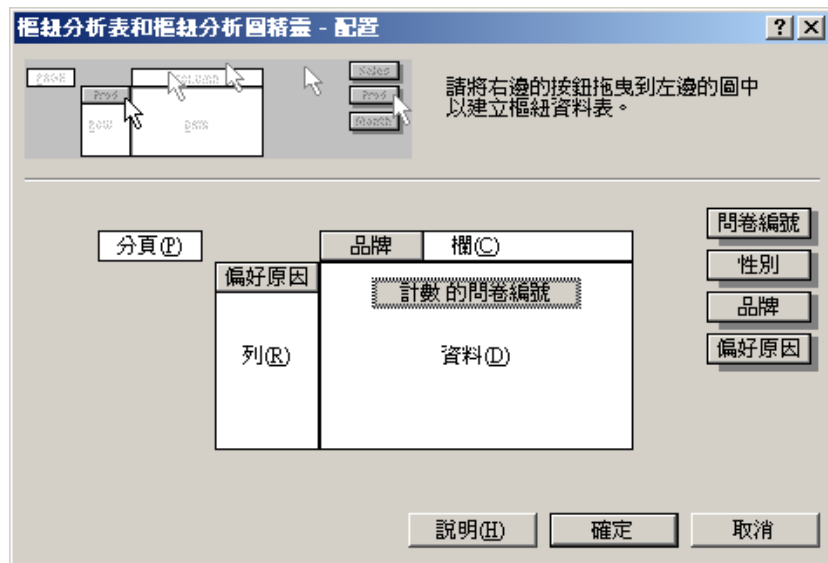
- 練習：針對前例『問卷資料』，求性別交叉使用品牌之交叉分析表，並以卡方檢定進行分析兩者是否存有關聯性。
($\alpha=0.05$)
- 由M15之P值為 $0.001 < \alpha = 0.05$ ，可知性別與品牌間存有高度之關聯性，使用品牌會隨性別而不同。
- 【問卷】工作表

L1		性別交叉品牌			
	L	M	N	O	P
1	性別交叉品牌	性別			
2	品牌	男	女	總計	
3	A牌	15	4	19	
4	B牌	5	12	17	
5	C牌	7	30	37	
6	總計	27	46	73	
7					

M15		=CHITEST(M3:N5,M11:N13)				
	L	M	N	O	P	Q
9	期望值	性別				
10	品牌	男	女			
11	A牌	32.0	1.3			
12	B牌	4.0	13.4			
13	C牌	3.6	38.6			
14						
15	卡方之P值	0.00				
16						

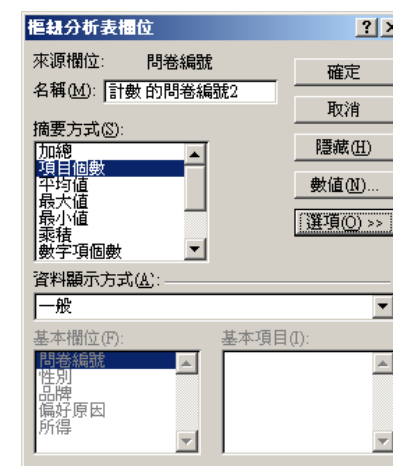
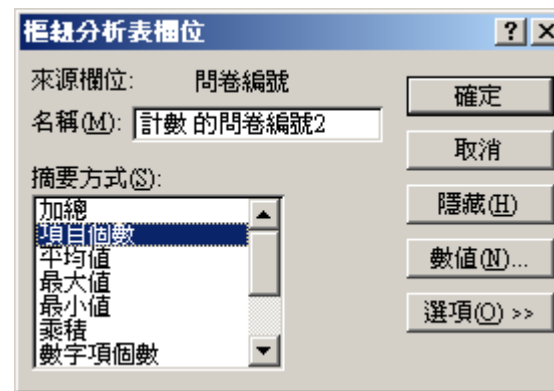
欄百分比

- 欲於原建立『品牌交叉偏好原因』之交叉表時，就一併求其人數及縱向百分比。
- 其操作步驟為：
 - 依前述操作步驟，操作到步驟9。
 - 『資料(D)』處所求算之統計量已改為計數
 - 再一次將〔問卷編號〕拖曳到『資料(D)』處，使『資料(D)』處擁有兩個內容
- 【欄百分比】工作表



欄百分比

- 於「加總的問卷編號」上雙按，仍將其『摘要方式(S):』改為「項目個數」，以求算出現次數
- 按「選項」鈕
- 按『資料顯示方式(A):』處之向下箭號，選取使用「總欄數的百分比」
- 按兩次「確定」鈕，回『樞紐分析表和樞紐分析圖精靈-步驟3之3』對話方塊

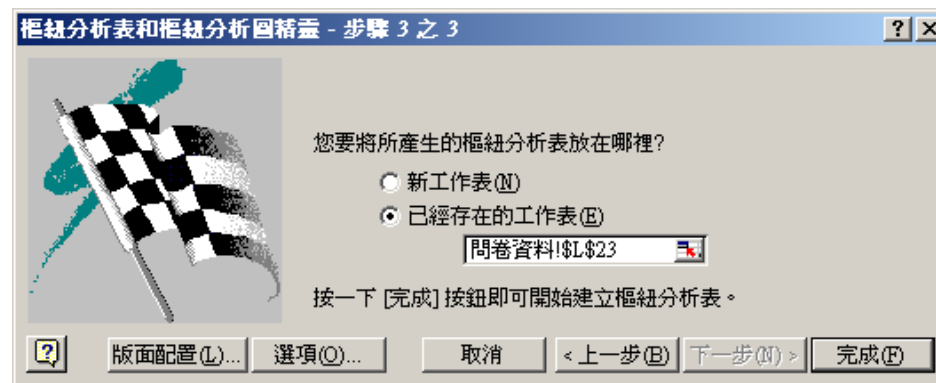


欄百分比

- 先選「已經存在的工作表(E)」項，續選按L23儲存格，擬將交叉分析表安排於目前工作表之L23
- 完成設定，按〔完成〕離開
- 於L23處獲致含人數及縱向百分比交叉分析表
 - M25改為輸入"人數"，
 - M26改為輸入"% "，讓其標題之說明字串更貼切一點
 - 於N24:P24輸入品牌名稱
 - 於L25、L27與L29輸入偏好原因
 - 於L22輸入『品牌交叉偏好原因』文字

L23						
	L	M	N	O	P	Q
23			品牌			
24	偏好原因	資料	1	2	3	總計
25	1 計數的問卷編號		11	1	6	18
26	計數的問卷編號2		64.71%	9.09%	27.27%	36.00%
27	2 計數的問卷編號		2	8	5	15
28	計數的問卷編號2		11.76%	72.73%	22.73%	30.00%
29	3 計數的問卷編號		4	2	11	17
30	計數的問卷編號2		23.53%	18.18%	50.00%	34.00%
31	計數的問卷編號的加總		17	11	22	50
32	計數的問卷編號2的加總		100.00%	100.00%	100.00%	100.00%
33						

L29 3. 外型美觀						
	L	M	N	O	P	Q
22	品牌交叉偏好原因					
23			品牌			
24	偏好原因	資料	A牌	B牌	C牌	總計
25	1. 價格便宜	人數	11	1	6	18
26		%	64.71%	9.09%	27.27%	36.00%
27	2. 品質優良	人數	2	8	5	15
28		%	11.76%	72.73%	22.73%	30.00%
29	3. 外型美觀	人數	4	2	11	17
30		%	23.53%	18.18%	50.00%	34.00%
31	人數的加總		17	11	22	50
32	%的加總		100.00%	100.00%	100.00%	100.00%
33						
34						



欄百分比

- 由表上資料可看出，整體上消費者偏好其使用之品牌的主要原因，依序為『價格便宜』（36%）、『外型美觀』（34%）與『品質優良』（30%）。
- 另由前節之卡方檢定之結果，可知各品牌與其使用者偏好原因間，存有顯著之關聯性。
 - A牌之使用者，主要是因『價格便宜』（64.7%）而使用A牌產品。
 - B牌之使用者，主要是因『品質優良』（72.7%）而使用B牌產品。
 - C牌之使用者，主要是因『外型美觀』（50.0%）而使用C牌產品。

L23		fx					
	L	M	N	O	P	Q	R
23			品牌				
24	偏好原因	資料	1	2	3	總計	
25	1	計數 的問卷編號	11	1	6	18	
26		計數 的問卷編號2	64.71%	9.09%	27.27%	36.00%	
27	2	計數 的問卷編號	2	8	5	15	
28		計數 的問卷編號2	11.76%	72.73%	22.73%	30.00%	
29	3	計數 的問卷編號	4	2	11	17	
30		計數 的問卷編號2	23.53%	18.18%	50.00%	34.00%	
31	計數 的問卷編號 的加總		17	11	22	50	
32	計數 的問卷編號2 的加總		100.00%	100.00%	100.00%	100.00%	
33							

欄百分比

- 練習：
 - 針對前例『問卷資料』，求性別交叉使用品牌之交叉分析表，於表中同時顯示人數及其縱向之欄百分比，並解釋其結果
- 【問卷】工作表

	L	M	N	O	P	Q	R
1	性別交叉品牌						
2			性別 ▾				
3	品牌 ▾	資料 ▾	男	女	總計		
4	A牌	人數	15	4	19		
5		%	55.6%	8.7%	34.0%		
6	B牌	人數	5	12	17		
7		%	18.5%	26.1%	22.0%		
8	C牌	人數	7	30	37		
9		%	25.9%	65.2%	44.0%		
10	人數 的加總		27	46	73		
11	% 的加總		100.0%	100.0%	100.0%		
12							
13	整體言，使用C牌者最多(44.0%)，其次為A牌(34.0%)						
14	男性以使用A牌者居多(55.6%)						
15	女性以使用C牌者居多(65.2%)						
16							

區間分組

- 性別交叉所得之結果，幾乎是一種所得即獨立存在產生一系列內容，於資料分析時並無多大作用。
- 較理想之方式為將所得分組，以縮減其組數
- 【[未分組](#)】工作表

L3		25000			
	L	M	N	O	P
1	計數的性別	性別			
2	所得	1	2	總計	
3	25000		1	1	
4	26000	3	1	4	
5	26400	1		1	
6	27000	1		1	
7	28000	1		1	
8	28500	3		3	
9	30000	1	2	3	
10	30200	1		1	
11	31000	1		1	
12	32000		1	1	
13	32900	1		1	
14	35600	1	1	2	
15	36000	1	1	2	
16	36400		1	1	

區間分組

- 若要將所得分為『未滿四萬』與『四萬及以上』兩組。
- 於資料表尾部，新增一欄
 - =IF(E2<40000,"1)未滿四萬","2)四萬及以上")簡化成兩組
- 前加數字，是為方便樞紐分析表依序排列分組內容；
- 若無，將會出現『四萬及以上』反排於『未滿四萬』之前
- 重建一次樞紐分析表，即可得到經縮減組數後之交叉表。
- 【[分組](#)】工作表

F2		=IF(E2<40000,"1)未滿四萬","2)四萬及以上")							
	D	E	F	G	H	I	J	K	L
1	偏好原因	所得	所得分組		計數 的所得分組	性別			
2	1	28000	1)未滿四萬		所得分組	1	2	總計	
3	2	30000	1)未滿四萬		1)未滿四萬	17	10	27	
4	1	26000	1)未滿四萬		2)四萬及以上	10	13	23	
5	2	32000	1)未滿四萬		總計	27	23	50	
6	2	45000	2)四萬及以上						
7	3	54000	2)四萬及以上						
8	1	31000	1)未滿四萬						
9	3	62000	2)四萬及以上						

區間分組

- 求出人數及縱向之欄百分比：

M1		fx				
	M	N	O	P	Q	R
1			性別			
2	所得分組	資料	男	女	總計	
3	1)未滿四萬	人數	17	10	27	
4		%	63.0%	43.5%	54.0%	
5	2)四萬及以上	人數	10	13	23	
6		%	37.0%	56.5%	46.0%	
7	人數的加總		27	23	50	
8	%的加總		100.0%	100.0%	100.0%	
9						

區間分組

- 練習：針對『問卷資料』，將所得分為
 - 『1.三萬以下』
 - 『2.三萬~五萬』
 - 『3.五萬及以上』
- 使用品牌交叉分組所得之交叉分析表，於表中同時顯示人數及其縱向之欄百分比，並解釋其結果。

	I	J	K	L	M	N	O
2	品牌交叉所得						
3			品牌 ▾				
4	所得分組 ▾	資料 ▾	A牌	B牌	C牌	總計	
5	1.三萬以下	人數	11			11	
6		%	64.7%	0.0%	0.0%	22.0%	
7	2.三萬~五萬	人數	6	10	7	23	
8		%	35.3%	90.9%	31.8%	46.0%	
9	3.五萬及以上	人數		1	15	16	
10		%	0.0%	9.1%	68.2%	32.0%	
11	人數 的加總		17	11	22	50	
12	% 的加總		100.0%	100.0%	100.0%	100.0%	
13							

F分配FDIST()

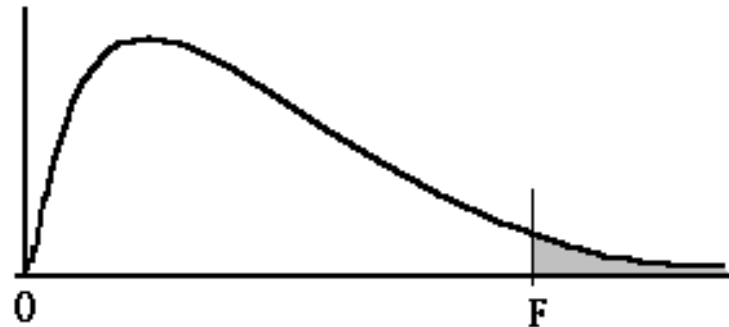
- **FDIST(F,分子的自由度,分母的自由度)**
- **F**為用來求算此函數的**F**值。由於**F**值是兩個均方相除：
- 故其自由度有兩個，一為分子的自由度；另一個為分母的自由度。
- 且因分子分母均為正值（均方），故其分配僅在**0**之右側而已。

$$F = \frac{MSA}{MSE} = \frac{\text{處理的均方}}{\text{誤差的均方}}$$

F分配FDIST()

- 本函數在求：於某兩個自由度下之F分配中，求自右尾累計到F值的總面積（機率）。即傳回F分配之右尾累計機率值（圖之陰影部份）：
- F分配之圖形及機率值，將隨自由度不同而略有不同。
- 以自由度為(2,10)與(3,15)之情況下，不同F值所求得之右尾累計機率分別為：
- **【FDIST】** 工作表

F6		fx		=FDIST(E6,3,15)					
	A	B	C	D	E	F	G	H	I
1	自由度為2,10				自由度為3,15				
2									
3	F值	右尾機率			F值	右尾機率			
4	4.10	0.050	← =FDIST(A4,2,10)		3.29	0.050	← =FDIST(E4,3,15)		
5	5.46	0.025	← =FDIST(A5,2,10)		4.15	0.025	← =FDIST(E5,3,15)		
6	7.56	0.010	← =FDIST(A6,2,10)		5.42	0.010	← =FDIST(E6,3,15)		
7									



$$F = \frac{MSA}{MSE} = \frac{\text{處理的均方}}{\text{誤差的均方}}$$

F分配反函數FINV()

- **FINV**(左尾機率,分子的自由度,分母的自由度)
- 本函數用以於已知自由度之**F**分配中，求某累計機率所對應之**F**值。
- 由於**F**分配之圖形及機率值，將隨自由度不同而略有不同
- 表是以自由度為**(2,10)**之情況下，所求得之結果。
- **【FINV】** 工作表

F6		fx		=FINV(E6,2,10)				
	A	B	C	D	E	F	G	H
1	自由度為2,10				自由度為2,10			
2								
3	F值	右尾機率			右尾機率	F值		
4	4.10	0.050	← =FDIST(A4,2,10)		0.050	4.10	← =FINV(E4,2,10)	
5	5.46	0.025	← =FDIST(A5,2,10)		0.025	5.46	← =FINV(E5,2,10)	
6	7.56	0.010	← =FDIST(A6,2,10)		0.010	7.56	← =FINV(E6,2,10)	
7								

F分配反函數FINV()

- 練習：
- 查兩個自由度（d.f.）分別為1~10之情況下，單尾機率為5%之F值
- 【FINV1】工作表

	B3		f _k	=FINV(\$B\$1,B\$2,\$A\$3)							
	A	B	C	D	E	F	G	H	I	J	K
1	α=	0.05									
2	分子 d.f.										
3	分母 d.f.	1	2	3	4	5	6	7	8	9	10
4	1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
5	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
6	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
7	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
8	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
9	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
10	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
11	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
12	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
13	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98

兩常態母體之變異數檢定

- FTEST(範圍1,範圍2)
- 可傳回兩組資料（樣本數允許不同）變異數是否存有顯著差異的F檢定之右尾機率值（P值）。
- 判斷檢定結果時很簡單，只須看此P值之二分之一是否小於所指定顯著水準之 α 值。

兩常態母體之變異數檢定

- 本函數可用來測試兩組樣本的變異數是否相同？即變異數同質性的檢定，其虛無假設與對立假設分別為：
 - $H_0: \sigma_1^2 = \sigma_2^2$ （兩變異數相等）
 - $H_1: \sigma_1^2 \neq \sigma_2^2$ （兩變異數不等）
- 要檢定甲乙兩班之母體變異數是否相同（ $\alpha=0.05$ ）？隨機抽得下示資料，以
 - $=FTEST(B2:B10, C2:C11)/2$
 - 求得其右尾機率（P值）並將其除以2，其值為 $0.26 > \alpha=0.05$ ，故無法棄卻兩變異數相等之虛無假設。
- **【FTEST】** 工作表

B15		=FTEST(B2:B10,C2:C11)/2				
	A	B	C	D	E	F
1		甲班	乙班			
2		80	85			
3		75	45			
4		35	69			
5		66	72			
6		84	80			
7		83	91			
8		58	68			
9		75	85			
10		92	87			
11			86			
12						
13	變異數	293.50	189.73			
14						
15	FTEST	0.26	<-- =FTEST(B2:B10,C2:C11)/2			
16						

先檢定變異數再進行均數檢定

- 當以t檢定，進行兩獨立樣本均數檢定時，將視其變異數相同或不同而使用不同之計算方法。
- 進行兩班抽樣成績之均數檢定：
 - 假設變異數相等（或不等）後，才來進行t檢定。
 - 這種假設合理否？誰都不知道！
 - 所以，先以F檢定，判斷其變異數是否相同
- 【成績】工作表

C13		fx =VAR(C2:C10)				
	A	B	C	D	E	F
1		甲班	乙班			
2		92	85			
3		75	64			
4		58	69			
5		66	81			
6		84	88			
7		78	91			
8		58	68			
9		75	75			
10		90	87			
11			91			
12						
13	變異數	157.86	98.75			
14	均數	75.11	79.90			
15						

先檢定變異數再進行均數檢定

- 由其F10處之P值為 $0.27 > \alpha = 0.05$ ，故無法棄卻兩變異數相等之虛無假設。
- 故可使用『兩獨立樣本均數檢定（變異數相同）』之方法進行檢定。
- 假定，要判斷在 $\alpha = 0.05$ 之顯著水準下，乙班之平均成績是否高過甲班？
- 由於是變異數相同，t檢定之類型為2。
- 且虛無假設與對立假設分別為：
 - $H_0: \mu_1 = \mu_2$
 - $H_1: \mu_1 < \mu_2$
 - 故此類檢定為單尾檢定

	A	B	C	D	E	F	G	H
1		甲班	乙班					
2		92	85		F 檢定：兩個常態母體變異數的檢定			
3		75	64					
4		58	69			甲班	乙班	
5		66	81		平均數	75.11111	79.9	
6		84	88		變異數	157.8611	102.9889	
7		78	91		觀察值個數	9	10	
8		58	68		自由度	8	9	
9		75	75		F	1.532797		
10		90	87		P(F<=f) 單尾	0.268464		
11			91		臨界值：單尾	3.229587		
12								
13	變異數	157.86	98.75					
14	均數	75.11	79.90					
15								

先檢定變異數再進行均數檢定

- **=TTEST(C2:C11,B2:B10,1,2)**
- 或『資料分析』之「t檢定：兩個母體平均數差異檢定」
- 假設變異數相等」均可進行檢定
- 無論由**B16**或**F26**之單尾P值來看，均顯示其值**0.19 > $\alpha=0.05$** ，故仍無法棄卻兩班之均數相等的虛無假設，所以並無法證明乙班之平均成績會高過甲班。

B16		=TTEST(C2:C11,B2:B10,1,2)							
	A	B	C	D	E	F	G	H	I
16	TTEST	0.19			t 檢定：兩個母體平均數差的檢定，假設變異數相等				
17									
18						甲班	乙班		
19					平均數	75.11111	79.9		
20					變異數	157.8611	102.9889		
21					觀察值個數	9	10		
22					Pooled 變異數	128.8111			
23					假設的均數差	0			
24					自由度	17			
25					t 統計	-0.92			
26					P(T<=t) 單尾	0.19			
27					臨界值：單尾	1.74			
28					P(T<=t) 雙尾	0.37			
29					臨界值：雙尾	2.11			
30									

t 檢定：兩個母體平均數差的檢定，假設變異數相等

輸入

變數 1 的範圍(1):

變數 2 的範圍(2):

假設的均數差(D):

☒ 標記(L)

α (A):

輸出選項

☒ 輸出範圍(O):

☐ 新工作表(P):

☐ 新活頁簿(W):

確定 取消 說明(H)

先檢定變異數再進行均數檢定

- 練習：
 - 以F檢定，判斷北區與南區給予剛畢業之大學生的薪資變異數是否相等（ $\alpha=0.05$ ）？
 - 續以適當之t檢定，判斷北區給剛畢業之大學生的平均薪資是否高過南區？
- 由其G10處之P值為 $0.1717 > \alpha=0.05$ ，故無法棄卻兩變異數相等之虛無假設，故得使用兩個母體變異數相等之均數檢定。
- 【薪資】工作表

	A2		f _x	
	A	B	C	D
1	剛畢業之大學生的薪資			
2		北區	南區	
3		28,500	30,000	
4		33,000	26,000	
5		30,000	25,000	
6		36,000	22,500	
7		29,000	31,000	
8		35,000	30,600	
9		29,800	20,500	
10		27,500	20,000	
11		30,500	26,000	
12				

	F	G	H	I
2	F 檢定：兩個常態母體變異數的檢定			
3				
4		南區	北區	
5	平均數	25733.33	31033.33	
6	變異數	17627500	8772500	
7	觀察值個數	9	9	
8	自由度	8	8	
9	F	2.009404		
10	P(F<=f) 單尾	0.1717		
11	臨界值：單尾	3.438103		
12				

先檢定變異數再進行均數檢定

- 由G26之單尾P
 - 值 $0.00348 < \alpha = 0.05$
 - 故得棄卻兩區均數相等之虛無假設
 - 所以北區給剛畢業之大學生的平均薪資高過南區

	F	G	H	I	J
16	t 檢定：兩個母體平均數差的檢定，假設變異數相等				
17					
18		北區	南區		
19	平均數	31033.33	25733.33		
20	變異數	8772500	17627500		
21	觀察值個數	9	9		
22	Pooled 變異數	13200000			
23	假設的均數差	0			
24	自由度	16			
25	t 統計	3.094533			
26	P(T<=t) 單尾	0.00348			
27	臨界值：單尾	1.745884			
28	P(T<=t) 雙尾	0.006959			
29	臨界值：雙尾	2.119905			
30					

單因子變異數分析

- 變異數分析的另一種用途，是用來檢定多組（>2）母群平均數是否相等。
- 其虛無假設與對立假設為：
 - $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ （每組之均數相等）
 - H_1 : 至少有兩個平均數不相等
- 假定資料為調查各地區對政府施政的整體滿意程度（滿分為100分），試以 $\alpha=0.01$ 之顯著水準，檢定各地區之滿意程度是否存有顯著差異？

	A	B	C	D	E
1	對政府的整體滿意度(滿分為100)				
2		北區	中區	南區	
3		75	65	85	
4		62	70	86	
5		50	80	87	
6		58	85	90	
7		82	82	95	
8		81	78	88	
9		75	85	75	
10		45	90	86	
11		65	84		
12		70			
13					

單因子變異數分析

- 將『分組方式』安排為「逐欄(C)」
- 點選「類別軸標記在第一列上(L)」(因各組資料均含標題之字串標記)
- α 設定為0.01
- 設定輸出範圍：於目前工作表F2
- 按〔確定〕鈕，即可獲致單因子變異數分析之ANOVA表
- 依此結果：自由度為(2,24)，F值10.80>臨界值5.61（K13處之P值0.00< α =0.01），故可知三個地區之整體滿意度存有顯著差異。南區的滿意度要比其餘兩區來得高。

	A	B	C	D	E	F	G	H	I	J	K	L
1	不同廣告方式所獲得之回應人數											
2						單因子變異數分析						
3	全版	半版	1/4版	小廣告								
4	1250	1083	850	660		摘要						
5	1324	1400	755	605		組	個數	總和	平均	變異數		
6	1600	1385	623	580		全版	6	7041	1173.5	73407.9		
7	890	680	600	856		半版	5	5416	1083.2	100081.7		
8	926	868	701	964		1/4版	7	5071	724.4286	7933.619		
9	1051		782			小廣告	5	3665	733	28403		
10			760									
11												
12						ANOVA						
13						變源	SS	自由度	MS	F	P-值	臨界值
14						組間	961279.6	3	320426.5	6.556359	0.003148	5.010287
15						組內	928580	19	48872.63			
16												
17						總和	1889860	22				

單因子變異數分析

輸入
 輸入範圍(I):

分組方式: ☒ 逐欄(C) ☐ 逐列(R)

☒ 類別軸標記是在第一列上(L)

α (A):

輸出選項
☒ 輸出範圍(O):

☐ 新工作表(P):

☐ 新活頁簿(W)

單因子變異數分析

• 練習：

- 某公司於報紙上進行廣告，表為不同方式廣告當天所獲得之回應人數。
- 試以 $\alpha=0.05$ 之顯著水準，檢定不同方式廣告之回應人數是否存有顯著差異？
- $F=6.56$ ， $d.f.=3,19$ ， $P=0.00$ ，不同方式廣告之回應人數間存有顯著差異。
- 全版與半版廣告之回應人數高於1/4版與小廣告

• 【廣告】工作表

	A	B	C	D	E
1	不同廣告方式所獲得之回應人數				
2					
3	全版	半版	1/4版	小廣告	
4	1250	1083	850	660	
5	1324	1400	755	605	
6	1600	1385	623	580	
7	890	680	600	856	
8	926	868	701	964	
9	1051		782		
10			760		
11					

	H	I	J	K	L	M
3	單因子變異數分析					
4						
5	摘要					
6	組	個數	總和	平均	變異數	
7	全版	6	7041	1173.5	73407.9	
8	半版	5	5416	1083.2	100081.7	
9	1/4版	7	5071	724.4	7933.6	
10	小廣告	5	3665	733.0	28403.0	
11						

	H	I	J	K	L	M	N
13	ANOVA						
14	變源	SS	自由度	MS	F	P-值	臨界值
15	組間	961279.6	3	320426.55	6.56	0.00	3.13
16	組內	928580.0	19	48872.63			
17							
18	總和	1889859.7	22				
19							