

查、某一都市研究，在臺南地區作隨機抽樣調查，發現南市居民在「環境品質」與「自有住宅房屋市價」這兩項變數之間有以下關係：

	環境品質評估				
	很滿意	有點滿意	沒意見	有點不滿	很不滿意
自有住宅房屋市價 (單位：元)					
200 萬以下	0	6	15	10	4
200~299 萬	1	7	20	6	1
300~399 萬	10	0	18	7	0
400 萬以上	0	5	15	15	0
計 (總計 140)	11	18	68	38	5

- 問題 (一)：根據這個交叉表，計算得到  $\chi^2 = 47.00$  ( $p=0.000$ )，請問這發現有何意義？  
 (二)：請問由這個交叉表發現，可否推測本樣本代表母群體？  
 (三)：有人說以上所計算的這個交叉表的  $\chi^2$  值有明顯的錯誤，請問為甚麼？  
 (四)：上述錯誤應如何改正？  
 (五)：改正後的  $\chi^2$  值是否顯示本樣本可代表母群體？ (每小題 4%；共 20%)

貳、某城市為了更新一塊大街廊舊商業區，擬在「區段撥收」、「市地重劃」、「聯合開發」與「居民自辦重劃」四種方式中，選擇一種來進行土地取得。經由民意調查居民偏好，得到以下反應：

地區	土地取得方式			
	區段撥收	市地重劃	聯合開發	居民自辦重劃
街廓東區	48	38	42	53
街廓西區	39	43	50	54
街廓南區	42	50	47	44
街廓北區	46	48	46	52

再經由拉丁方格分析，得到以下變異量分析結果：

source of variation	degree of freedom	sum of square	mean square	F
rows	3	17.25	5.75	0.9
columns	3	114.75	38.25	5.9
treatments	3	174.75	58.25	9.0
error	6	39.00	6.5	
total	15	345.75		

- 問題 (一)：請問這四種土地取得方式所具影響效果是否有所差異？由甚麼證明？  
 (二)：請問這四個街廓內區位所具影響效果是否有所不同？由甚麼證明？  
 (三)：請問這四種土地取得方式，對於四街廓區位間的居民選擇，是否各具影響效果？由甚麼證明？

提示： $F_{0.05}(d.f., 3) = 4.76$  (每小題 5%；共 15%)

李、某研究依照 Shevky & Bell 的「社會區域分析」(SOCIAL AREA ANALYSIS) 架構，企圖以六個變數建構較少個的指標，來將 474 個都市戶籍區域作歸類。其統計分析結果報表如下：

Analysis number 1 Listwise deletion of cases with missing values

	Mean	Std Dev	Label
AGE	37.18614	11.78724	平均戶長年齡
EDLEVEL	13.49156	2.88485	平均戶長教育年數
JOB CAT	2.05907	1.40548	主要戶長職業
EXPENSE	6806.43460	3148.25528	平均家戶每年開銷(百元)
INCOME	13767.82700	6830.26459	平均家戶戶長每年所得(百元)
ETHNIC	81.10970	10.06094	地區每百戶家戶中本省籍家戶數

Number of Cases = 474

Initial Statistics:

Variable	Communality	* Factor	Eigenvalue	Pct of Var	Cum Pct
AGE	1.00000	1	3.14858	52.5	52.5
EDLEVEL	1.00000	2	1.06553	17.8	70.2
JOB CAT	1.00000	3	1.00128	16.7	86.9
EXPENSE	1.00000	4	.45494	7.6	94.5
INCOME	1.00000	5	.22785	3.8	98.3
ETHNIC	1.00000	6	.10242	1.7	100.0

Factor Matrix:

	Factor 1	Factor 2	Factor 3
AGE	-.20536	.83788	-.48681
EDLEVEL	.78905	-.17286	.20238
JOB CAT	.89807	.07621	-.18216
EXPENSE	.92781	.15747	-.14951
INCOME	.94165	.09178	.02703
ETHNIC	.01923	.54281	.83389

Final Statistics:

Variable	Communality	* Factor	Eigenvalue	Pct of Var	Cum Pct
AGE	.95290	1	3.14858	52.5	52.5
EDLEVEL	.69337	2	1.06553	17.8	70.2
JOB CAT	.77528	3	1.00128	16.7	86.9
EXPENSE	.90780				
INCOME	.89585				
ETHNIC	.99039				

- 問題(一)：請問這個研究所用的是甚麼統計方法(請寫出名稱)？  
 (二)：根據這個統計分析報表發現的那一部分，可以建構成三個「社會區域」指標來對 474 個戶籍地區作分類？  
 (三)：如果以三個「社會區域」指標來區分這 474 個戶籍區域，約可準確區分的百分比有多少？從那個數據來判斷之？  
 (四)：如果要將這三個「社會區域」指標命名，將要依據統計報表的那一部分發現來從容之？

(每題 5 分；共 20 分)

肆、某君欲探討都市居民教育水準之特性及差異問題，選取了台灣地區之各級大小都市 30 餘個，從事複迴歸分析，應變數設定為「都市中大專及以上教育程度者佔總人口之比率」，試回答：

1. 某君應如何決定其所應使用之解釋變數?(4%)
2. 某君不確定應使用那些變數，因而以許多可能與教育程度有關之變數從事 stepwise regression 分析，來找出影響程度最高之幾個變數加入其迴歸模型中，你對他的做法有何意見?(5%)
3. 從事迴歸分析時，當你拿到電腦的報表時，請說明你如何分析及研判你的迴歸模型是否適當、有何問題?(5%)
4. 某君擬了二個迴歸模型，模型 A 之  $R^2=0.94$ ，模型 B 之  $R^2=0.89$ ，A 之  $R^2$  較高，故某君選擇 A 為其所建立之迴歸模型來分析教育水準問題。你有何意見?(5%)
5. 某君之模型中有人口數之解釋變數，並發現人口數越多的都市，教育水準越高。某乙不同意，認為係教育水準較高者湧入大都市，才造成其人口之增加。就迴歸模型來說，上述爭議表示些什麼問題?(5%)
6. 就一迴歸模型而言，假如遺漏了一個應該在內的解釋變數其後果如何?假如加入了一個不應加入的解釋變數其後果如何?(5%)

(共 29%)

伍、某君從事一產品之市場調查，隨機抽取顧客  $n$  為樣本，詢問其是否喜歡此產品，在 98% 之信賴度下，希望抽樣誤差  $\leq 0.05$ ，試問：

1. 初步估計母體中有 35% 的顧客會喜歡此產品，此時  $n$  應至少為多少?(8%)
2. 假如沒有任何資訊可估計母體中有多少顧客會喜歡此產品時， $n$  應至少為多少?(8%)

(共 16%)