

樣本數之決定及抽樣方法

在醫藥衛生的研究中，樣本數的計算常常會被忽略，在較具規模的研究補助機構如國科會中，提報研究計畫時，樣本數如何算出就必須清楚說明，絕不能含混帶過。對多數研究者來說，如何估算樣本數？一直是個困難的問題，很多研究結果失敗並非研究設計或介入措施不好，而是研究計畫之初沒有估算樣本數，樣本數太少往往也是造成研究結果達不到顯著差異的原因之一，因此，如何估算樣本數？樣本數決定後如何抽樣？如何使用軟體快而有效率的達到抽樣結果是本章要介紹的重點。

■ 為什麼需要估計樣本數？

樣本數會影響統計估計的準確度，例如：某一婦產科病房想改建為坐月子中心，應先調查該醫院轄區社區內一年內生產比率有多高，以作為考量該開幾床時之參考，但生產比率要如何估算？不可能挨家挨戶打電話訪問，此時便要做抽樣調查，若抽樣不好、估算準確度不高便會影響到營運。

$$SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

上列是平均值的標準差公式，若樣本之間的異質性很大，標準差會愈大，error 就會提高，所需的樣本數也會愈多。

此外估算樣本數還有一個目的，就是要去算統計檢力 (power)，樣本數太小無法達統計顯著差異。樣本數並非愈多愈好，過度抽樣是浪費資源，在資料處理時也沒有效率。

■ 虛無假設與對立假設

1. 虛無假設 (null hypothesis, H_0) ; $H_0 : \mu_1 = \mu_0$

所謂虛無假設是指假設所比較母群體間之母數“無差異”。亦即，要比較之兩組所對應出的母體平均數 (μ) 是相同的；有時兩組比較的不一定是平均數，也有可能是比率或次數等。

2. 對立假設 (alternative hypothesis, H_1) ; $H_1 : \mu_1 \neq \mu_0$

而對立假設是指兩組所對應出的母體平均數 (μ) 是不同 (有差異) 的。無論何種統計方法都會算出p值，而p值是指虛無假設成立的機會，若檢定有統計差異則推翻虛無假設，接受對立假設；若檢定無統計差異則不要推翻 (接受) 虛無假設。研究者真正所要的結果是對立假設成立，即實驗組不等於對照組。

■ 第一類誤差(α)與第二類誤差(β)

1. 第一類誤差(α)：

要推翻或接受無效假設都要由臨界值來判定，一般臨界值是訂 0.05，如果 p 值小於 0.05，可說無效假設成立的機會太低了，無法接受要推翻掉，當我們說推翻掉的同時，會有 5% 不正確的風險，故拒絕無效假設後所發生的不正確性即稱為第一類誤差。臨界值也可訂 0.01 或 0.001，一般來說，會對民眾生活造成影響的研究，例如探討石化區居民生活的研究，臨界值最好訂嚴謹些，以免研究結果造成居民恐慌。

2. 第二類誤差(β)：

若檢定結果在顯著水準5%下無統計顯著差異，因此不推翻虛無假設，意指“接受虛無假設有5%不正確的風險。”，稱為第二類誤差。第二類誤差研究者可自訂，通常是訂20%，而所謂的檢力 (power) 是1-第二類誤差，若研究者將第二類誤差訂為20%，其檢力就是80%。

■ 精確度 (precision) 與敏感度 (difference, d) (參考表一)

1. 精確度

事實上是沒有差異的，而研究者下的結論也是沒有差異，即稱為精確度。

2. 敏感度

事實上是具有差異，而研究者也正確的拒絕了無效假設，稱為敏感度。此敏感度也就是所謂的檢力 (power)，有些統計軟體顯示的不是 power 而是 d，就是在問研究者希望研究的兩個群體監所要檢定的依變項達到多敏感的程度可被偵測出來。舉例來說，某醫學中心壓瘡盛行率為千分之一，我們現在也要進行壓瘡研究，希望樣本之間差距千分之一就被偵測出來，要符合這樣條件需多少樣本數？ 敏感度愈小所需的樣本數愈大。

■ 計算樣本數的軟體通常需要以下幾個參數： α 、error、d (敏感度)、precision 及 N (母群體約有多大)。

■ 檢力分析

檢力是指決定需要多大的樣本數來偵測出特定程度的差異，當研究結果無“統計差異”時，需再注意研究是否有足夠的檢力來偵測“臨床差異”。一般在計劃階段就應決定檢力要多大；檢力過低之負面影響為：因樣本數太少致無法找到合理的差異。

表1

事實 \ 檢定結果	有差異	沒差異
拒絕 H_0 (有差異)	敏感度 (Power=1- β)	第一類 (α error)
接受 H_0 (沒差異)	第二類 (β error)	精確度

■ 樣本數估計原則

需回答下述四個問題

1. 臨界值為何 習慣上第一類誤差 α error=0.05

α error 是雙尾檢定，當 α error=0.05 時其對應的 z 值為 1.96。

當 α error=0.02 時其對應的 z 值為 2.33。

2. 檢力要多大？習慣上之檢力 $\text{power}=1-\beta=1-0.2=0.8$

β error 是單尾檢定，當 β error=20%，其對應的 z 值為 0.84。

當 β error=10%，其對應的 z 值為 1.28。

3. 要多大”臨床差異“來偵測差異存在？需由研究者判定

4. 研究變項的變異性為何？這會影響統計準確度，即影響所需樣本數

所謂的變異性轉成統計術語就是標準差 (σ)，此參數可由文獻或做預試而得。

計算樣本數的公式

依研究型態的不同，計算樣本數的公式也不同，以下幾例說明之：

$$n = \frac{Nz^2\sigma^2}{d^2(N-1) + z^2\sigma^2}$$

N ：母群體數

z ： z 值

d ：精確範圍

σ ：母群體變異數

■ 一般調查之樣本數 (平均值)

■ 一般調查之樣本數 (比率)

$$n = \frac{Nz^2p(1-p)}{d^2(N-1) + z^2p(1-p)}$$

N ：母群體數

z ： z 值

d ：精確範圍

p ：預期母群之盛行率

■ 比較一組樣本平均值

$$Z_\alpha = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad Z_\beta = \frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} \quad n = \left[\frac{(Z_\alpha - Z_\beta)\sigma}{\mu_1 - \mu_0} \right]^2$$

研究問題：脊柱側彎病人最大吸氣壓是否與常人相同？

假定：1. 常人最大吸氣壓 $\mu_0 = 110$ $\mu_1 = 20 \text{cmH}_2\text{O}$

2. 脊柱側彎病人最大吸氣壓臨床意義 $\leq 80 \text{cmH}_2\text{O}$

α ? Power? Significant difference? Variability?

$$n = \left[\frac{(Z_\alpha - Z_\beta)\sigma}{\mu_1 - \mu_0} \right]^2 = \left[\frac{(1.96 + 1.28)(20)}{110 - 80} \right]^2 = 4.67 \approx 5$$

■ 比較二組獨立樣本平均值

研究問題：高血壓患者血清中兒茶酚 (catecholamines) 與心跳的關係？

假定：1.高血壓患者血清中兒茶酚較高者的心跳顯著高於兒茶酚較低者

2.臨床顯著意義為每分鐘心跳差異大於10, 心跳標準差為每分鐘15下

α ? Power? Significant difference? Variability?

$$n = 2 \left[\frac{(Z_\alpha - Z_\beta)\sigma}{\mu_1 - \mu_0} \right]^2 = 2 \left[\frac{(1.96 + 0.84)(15)}{10} \right]^2 \approx 36$$

■ 比較一組樣本比率

研究問題：想瞭解寄發提醒函是否會增加體檢率，以前體檢率40%？

假定：1.改善25%即值得(從40%至50%體檢率)

2.接收第一類錯誤為2%及90%正確差異檢力

α ? Power? Significant difference? Variability?

$$n = \left[\frac{(Z_\alpha \sqrt{\pi_0(1-\pi_0)} - Z_\beta \sqrt{\pi_1(1-\pi_1)})}{\pi_1 - \pi_0} \right]^2$$

$$n = \left[\frac{(2.326\sqrt{0.4 \times 0.6} - (-1.28)(0.5 \times 0.5))}{0.4 - 0.5} \right]^2 = 317$$

■ 比較二組樣本比率

研究問題：想瞭解患者接受J5抗血清是否能有效預防革蘭式菌感染？

假定：1.感染率10%，臨床上意義需降至5%。

2.接收第一類錯誤為5%及90%正確差異檢力

$$n = \left[\frac{(Z_{\alpha} \sqrt{\pi'(1-\pi')} - Z_{\beta} \sqrt{\pi_t(1-\pi_t) + \pi_c(1-\pi_c)})}{\pi_c - \pi_t} \right]^2$$

α ? Power? Significant difference? Variability?

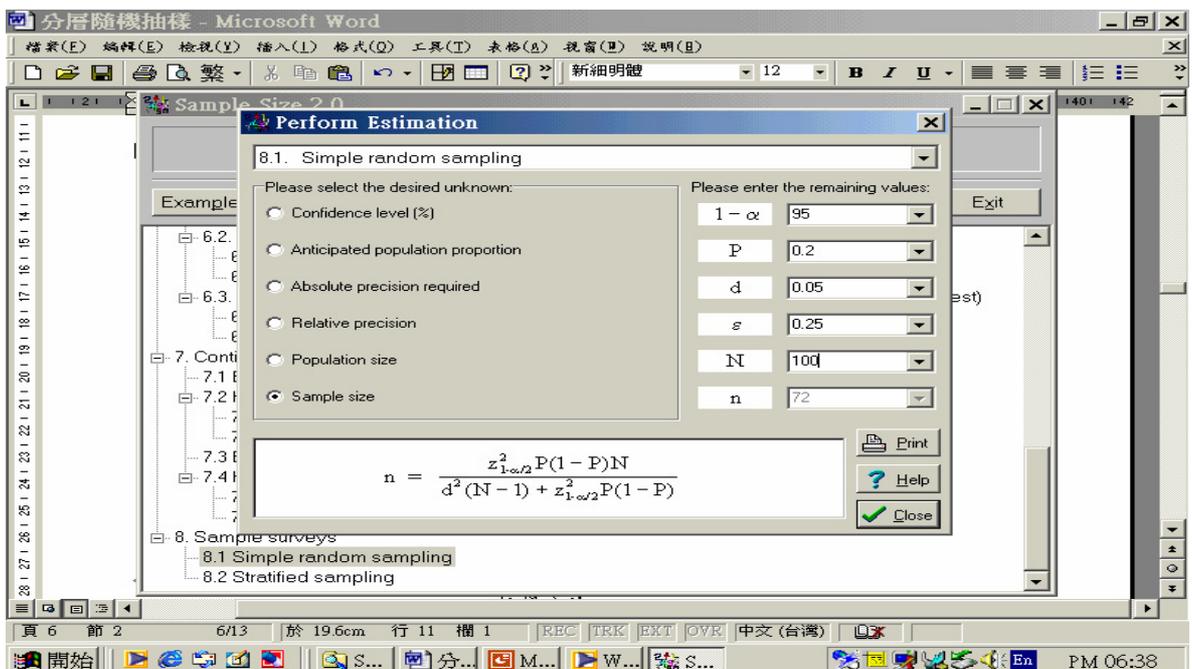
$$n = \left[\frac{(1.96 \sqrt{2 \times 0.075 \times 0.95} - (-1.28) \sqrt{(0.05 \times 0.95) + (0.1 \times 0.9)})}{0.05 - 0.10} \right]^2 = 5805$$

SSIZE 軟體估算樣本數之操作方法

用 SSIZE 軟體估算樣本數之前，需先界定研究型態、研究設計為何，以選擇正確的估算公式。但 SSIZE 軟體所列之研究設計選項中有些與實際研究情境有差異，例如：Hypothesis testing for two population means 是假設兩組情境相同，不同的是有無介入，未考慮到混淆因素，所以真正的樣本數，除了軟體計算結果外，還需視混淆變項有幾個，每多一個變項需再加 10 個樣本。

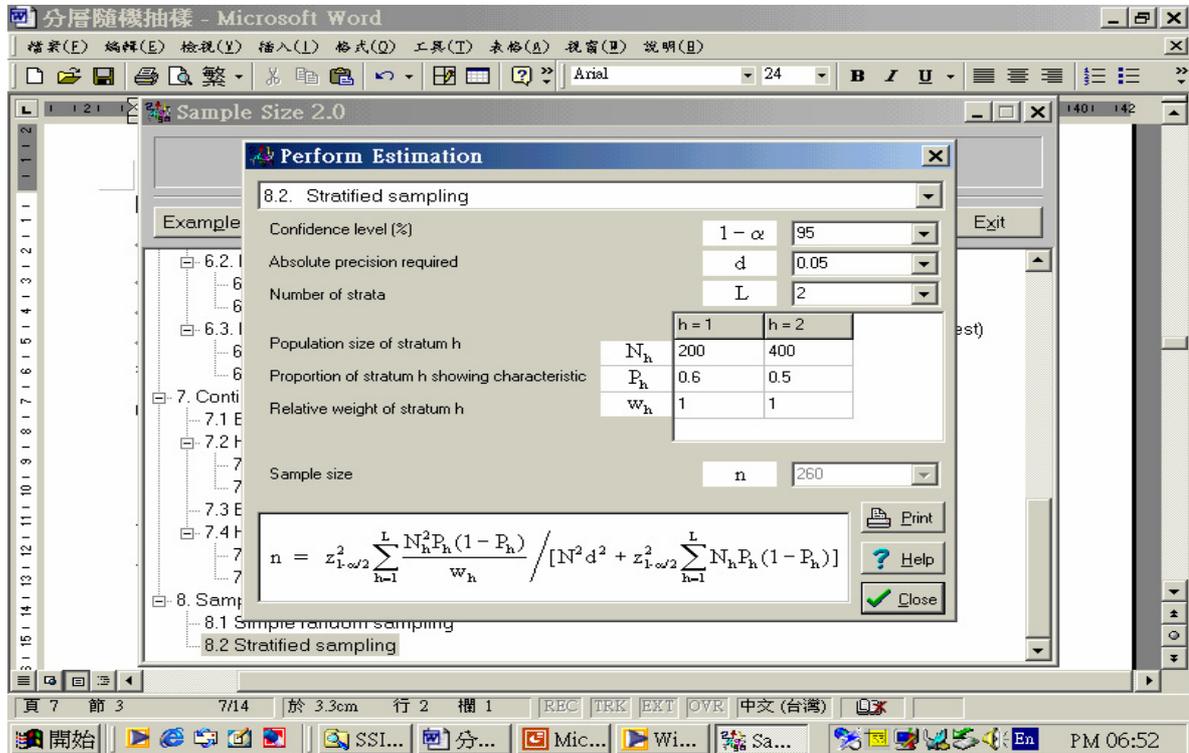
■ 以 8-1 Simple random sampling 為例

點選 Simple random sampling 後，再點選最上列之 Estimate，畫面如下：



依軟體所需之參數，輸入各項數值，即可算出樣本數，軟體所提供的公式若寫研究報告時應當要寫出。

■ 以8-2 Stratified sampling 為例，畫面開啟如下圖



結果所得的樣本數是每一層最少的人數，若有兩層還需將結果乘以 2。

抽樣方法

抽樣的方法可分為隨機抽樣(random sampling)與非隨機抽樣(non-random sampling)，而隨機抽樣方法又可細分有：簡單隨機抽樣、系統抽樣、分層抽樣、集束抽樣及多步驟隨機抽樣，以下就各種抽樣方法做說明。

■ 隨機抽樣(random sampling)

所謂的隨機抽樣是指母群體中每一個體皆有被抽中的機會，即其被抽中的機率應大於0；至於個體被抽中的機會應是隨機的，不加以人為安排。隨機抽樣的方式又可分為以下五種：

1. 簡單隨機抽樣 (simple random sample)：

每個體被抽中的機率是一樣的，適用於母群體不大或母群體大但每個個體均有編號的情境上，可以使用抽籤法或亂數表來達到抽樣目的。

抽籤法 - 先將母群體之所有個體統一編號，每一個體均編一號碼。

- 另將每一號碼書寫於同樣性質的卡片或球上。
- 將卡片或球充分混合後，再抽出卡片或球。
- 幸運抽獎、有獎問答即屬之。

亂數表 - 首先將母群體之所有個體統一編號，每一個體均編一號碼。

- 編號位數應與母群體數之位數相同，如母群體數為三位數，編號應為三位數，001, 002,..。
- 事先決定選取號碼之方向。
- 盲目隨手一指亂數表，以決定第一個數字。
- 若所得一組數字已大於母群體數或重複，則略過而選取下一組數字。

2. 系統抽樣 (systematic sampling) : 在母群體中每隔一定距離抽取一個樣本，若觀察的變項變化週期與抽樣距離相等時，雖然仍為隨機樣本，但其統計值與母數可能差距甚大而導致偏差。

舉例：某校每班 40 人，依身高由矮到高編號，欲得知該校平均身高，以 40 為組距，隨機號碼 2，即以每班第 2 號為樣本，所得樣本的平均值一定偏低；故進行系統抽樣時須留意抽樣距離與變化週期的關係。

3. 分層抽樣 (stratified sampling) : 在某些時候，除了要推估母群體外，對於次母群體也因有所瞭解，故以簡單或系統隨機抽樣抽取個次母群體中數個個體作為樣本，進行加權分析，藉由從次母群體所得到的個案推估母群體的狀況。

分層的原則：(1).層內個體應儘量相同，簡化母群體構造；(2).層間個體應儘量不同，以便包括母群體內各種特性；(3).等比率抽樣 (proportional sampling)各層中樣本數占全部樣本數的比率，應與每一層個體數占母群體的比率相同；(4).不等比率抽樣(disproportional sampling)應予事後加權。

舉例：欲調查某校高中學生對婚前性行為的意見，全校有 3600 人，其中高一有 1400 人，高二有 1200 人，高三有 1000 人，因不同年級對於婚前性行為的意見可能有所不同，故由各年級中分層抽取樣本；欲抽取 10% 的母群體作為樣本，則應由高一抽取 $1400 \times 10\% = 140$ 人，高二抽取 $1200 \times 10\% = 120$ 人，高三抽取 $1000 \times 10\% = 100$ 人當成樣本，再以簡單隨機或系統隨機抽樣由各年級抽取所需的人數，此種抽樣方式即分層抽樣。

4. 集束抽樣 (cluster sampling)：依研究目的先將母群體分為若干個團體，因團體內異質性高，但團體間的同質性高，故由母全體中抽取數個團體，以團體內所有的個體為研究的樣本，此即為集束抽樣。

步驟：(1)依某種標準 (研究目的)，將母群體分為若干類或組，每類分為稱為一個團體；(2)再於每一團體中以隨機取樣方式，抽取若干小團體，這些小團體中的各個體為樣本。

舉例：欲瞭解台北市大安區國小學童齲齒的盛行率，因位在同一區的學童生活狀況、口腔衛生習慣可能相類似，故從大安區所有的國小中選取兩間學校作為研究對象，以該校所有學童作為研究樣本進行齲齒盛行率的調查，用以推估台北市大安區國小學童齲齒的盛行率。

5. 多步驟隨機抽樣 (multistage random sampling)：當母群體甚大、人力不足時，分兩次或兩次以上抽樣，此即為多步驟隨機抽樣。

步驟：(1).依某種標準，將母群體分為若干層或組；(2).以隨機方式抽出幾層；(3).再從所抽取之層中，以隨機取樣抽出樣本加以分析；分兩次抽樣為兩段抽樣，一般超過三次時稱為多段取樣 (multiple-stage sampling)。

舉例：欲調查台灣省小學生學生之齲齒情形，先將台灣省 21 縣市依序編號，採簡單隨機抽樣的方式抽取 10 縣市，在抽得的 10 個縣市中再以簡單隨機抽樣於此 10 個縣市分別抽取 4 個鄉鎮，之後，以簡單隨機抽樣於所抽到的鄉鎮中再抽出一所小學，將所抽到的小學(共 $10 \times 4 \times 1 = 40$ 所學校)所有學生為視為樣本，調查此 40 所學校的齲齒情形，以推估台灣省小學學生的齲齒情形。

■ 非隨機抽樣(non-random sampling)

依研究者的需要或意願，選取具有某種特性之個體作為樣本，使得部分個體永遠沒有被抽中的機會，違反隨機的原則，即所謂的非隨機抽樣。如病例對照研究的病例組及對照組，經常是來自於醫院的病患，非隨機而來的；一般的街頭訪問，訪問的對象僅限於有前往到當地的民眾；藥物臨床試驗中，在醫療倫理考量為前提下，受試者都必須是自願的。

應用非隨機抽樣的原因有以下幾點：(1)母群體難以界定或接觸；(2)先趨研究；(3)個案稀少、不易獲取；(4)倫理或法律；(5)實務面考量。

研究者可依(1)依據專業判斷、母群體的特性及研究目的；(2)便利性：例如街頭調查、醫院的病人、某機構的員工，但需注意志願者效應 (volunteerism)；(3)依某種既定的標準來取樣，例如職業別、年齡別、教育程度、地區別；(4)經由少數個案之轉介，而得到其他個案等原則選擇非隨機抽樣方法。

運用 SPSS 軟體執行分層隨機抽樣

以護理人員所屬病房別（檔案 Nssample.xls）為例，要執行分層隨機抽樣，

步驟說明如下：

一、開啟 SPSS；叫出檔案 Nssample.xls

檔案叫出後看一下原始資料中個案分布情形

步驟：Analyze → Descriptive Statistics → Frequencies → Code 選入

結果：5400~5600 這幾個各案人數都只有個位數不適合分為一層，故 recode

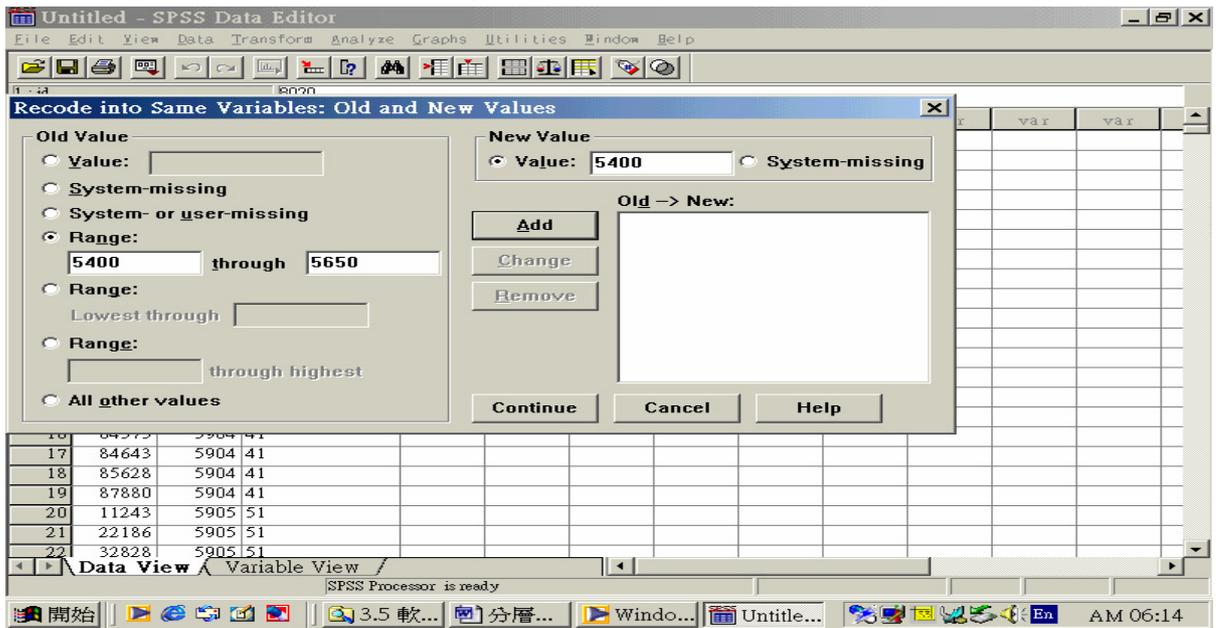
成一個單位。

二、RECODE（把個案少的單位合併，重新分層）

步驟：Transform → Recode → Into Same Variables → 將 Code 此變項

選入 → 在 Old and New Values 功能中將 5400~5600 此範圍之個

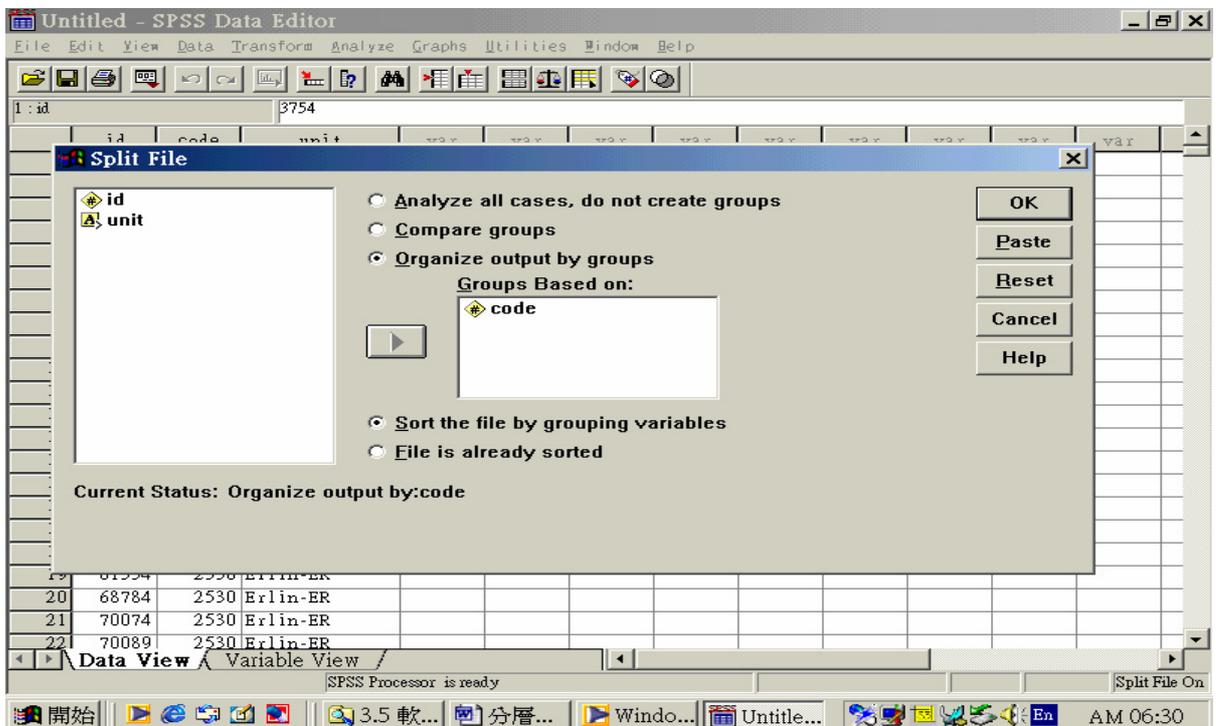
案 recode 成 5400；2560 之個案 recode 成 2500，如下圖



三、SPLIT (分割檔案或資料)

步驟：Data → Split File → 在 Organize output by groups 將 Code 此變項

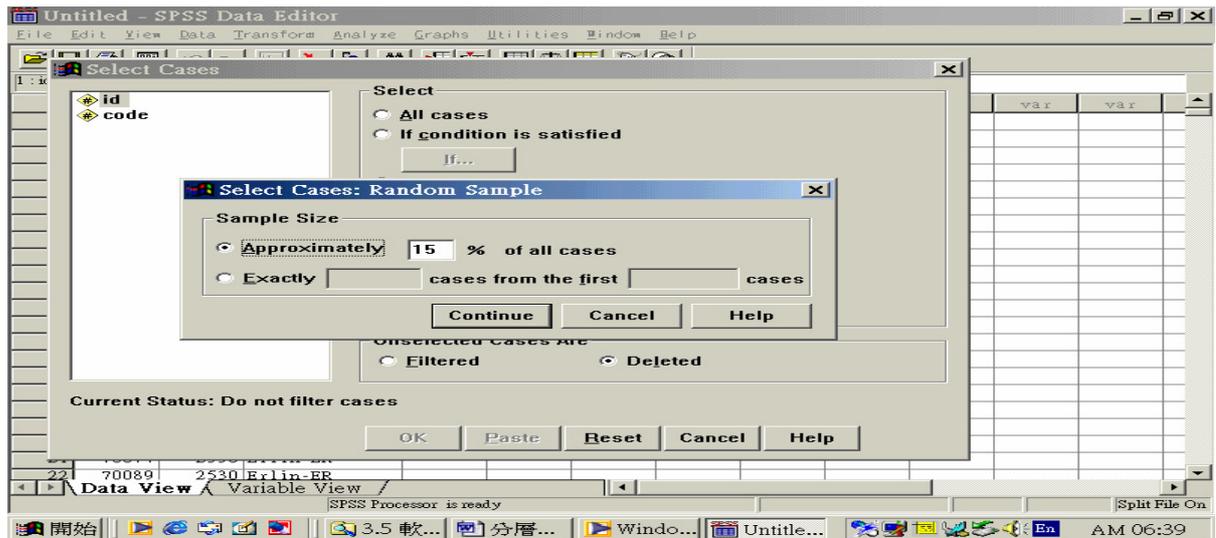
選入，畫面如下圖：



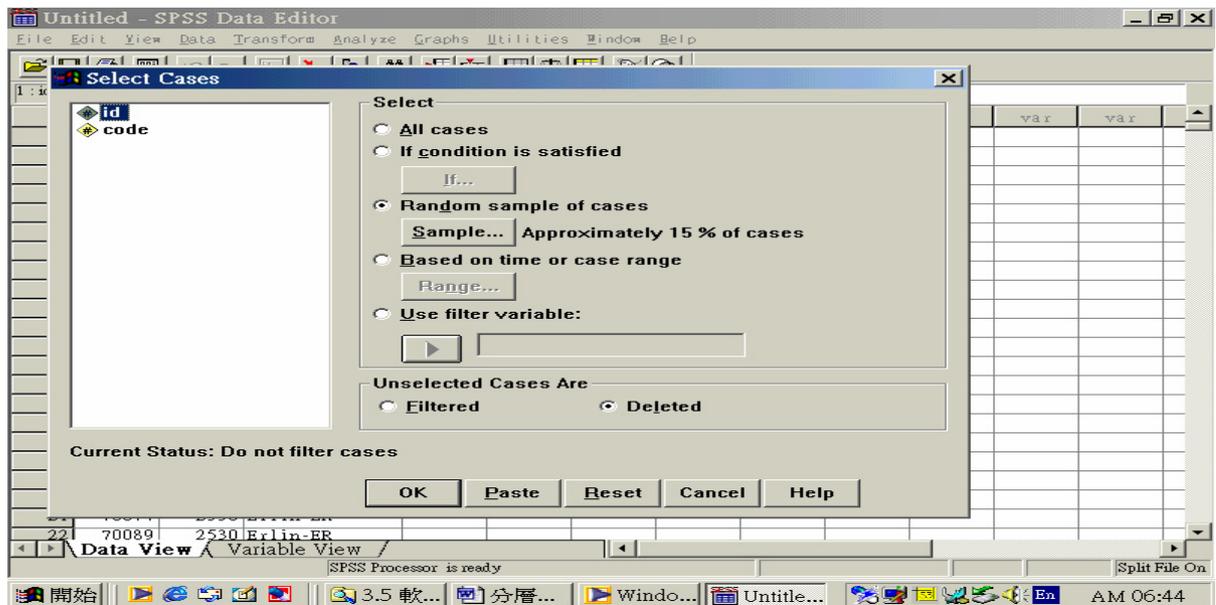
四、SELECT (選擇觀察值)

步驟: Data → Select cases → Random sample of cases → sample

中點入，畫面詢問 sample size 要由軟體估算，或研究者設定範圍，畫面如下：



Sample size 估算設定好後，回到select cases 畫面，在Unselected Cases Are...選deleted（未選到的個案刪除；Filtered是指未選到的個案畫斜線），畫面如下：



五、以上步驟依序執行後即完成分層隨機抽樣步驟，將最後結果存檔，存成 Excel 檔，才可將資料列應出來。