

壹、計畫介紹

計畫名稱：家庭動態調查：2000新抽樣本（RI2000）

調查執行期間：民國89年1-3月

抽樣使用母體資料年度：民國88年五歲年齡層的資料

母體定義：民國24年至民國43年出生之一般民眾

完成案數：1,958案

貳、加權說明

一、加權母體說明

此次加權為能反映訪問當年之人口結構，以民國 89 年內政部台閩地區人口統計資料為準，扣除外島人口，包括連江縣、金門縣、屏東縣琉球鄉、台東縣綠島鄉、台東縣蘭嶼鄉及澎湖縣等地區後，分別計算母體性別、年齡及地區層別的分佈。

在教育程度的母體參數部份，使用民國 89 年戶口普查資料，其方式為先扣除外籍人士(民國 89 年的戶口普查資料，利用「國籍代碼」變項挑出“本國人”(000)資料)，再將年齡限制為 46 歲以上 66 歲以下，並扣除上述外島人口。加權之變項、分組方式及母體參數如表一。

表一：加權之變項與分組方式

變項	樣本	母體		
性別	1. 男 2. 女	性別	人數	百分比
		1 男生	2,069,032	49.93
		2 女生	2,074,720	50.07
		全國人數 4,143,752 人		
年齡	1. 46-48 歲 2. 49-51 歲 3. 52-54 歲 4. 55-57 歲 5. 58-60 歲 6. 61-63 歲 7. 64-66 歲	年齡	人數	百分比
		1 46-48	957,317	23.10
		2 49-51	730,252	17.62
		3 52-54	616,720	14.88
		4 55-57	496,593	11.98
		5 58-60	481,150	11.61
		6 61-63	450,265	10.87
		7 64-66	411,455	9.93

變項	樣本	母體		
地區層別 ¹		地區層別	人數	百分比
	1. 新興鄉鎮	1. 新興鄉鎮	627,687	15.15
	3. 工商市鎮	3. 工商市鎮	573,823	13.85
	4. 綜合性市鎮	4. 綜合性市鎮	412,950	9.97
	5. 坡地鄉鎮	5. 坡地鄉鎮	240,865	5.81
	6. 山地及偏遠鄉鎮	6. 山地及偏遠鄉鎮	530,688	12.81
	7. 服務性鄉鎮	7. 服務性鄉鎮	431,876	10.42
	8. 省轄市	8. 省轄市	485,752	11.72
	9. 台北市	9. 台北市	549,850	13.27
	10. 高雄市	10. 高雄市	290,261	7.00
教育程度		教育程度	人數	百分比
	1. 無/自修/小學	1 小學及以下	2,391,953	57.76
	2. 初(國)中/初職	2 初(國)中/初職	530,854	12.82
	3. 高中普通科/高中職業科/高職	3 高中職	679,490	16.41
	4. 五專/二專/三專	4 專科	231,971	5.60
	5. 技術學院/大學/碩士/博士	5 大學及以上	307,103	7.42

二、加權方式說明

1. 不等機率加權

“不等機率抽樣加權權值”(Unequal probabilities of selection, sampling weights) 主要是要顯示某一受訪者在各分層中選的機率。

Step 1. 各分層裡每個人的原中選機率

$$f_{\text{selection}} = \text{PSU}_h * \text{SSU}_h * \text{TSU}_h = \frac{\alpha_h \beta_{ha}}{N_h} * \frac{b_h c_{ha\beta}}{\beta_{ha}} * \frac{c_h}{c_{ha\beta}} = \frac{\alpha_h b_h c_h}{N_h} \quad (1)$$

Step 2. 膨脹樣本後的個人中選機率

$$f_{(\text{county})} = \frac{\alpha_h b_h c_h * \text{膨脹倍數}_{(\text{county})}}{N_h} \quad (2)$$

Step 3. 加權權值的計算公式

$$w = \frac{1}{f_{(\text{county})}} \quad (3)$$

¹ 地區層別依據羅啟宏(1992)所著之「台灣省鄉鎮發展類型之研究」的分層原則將臺灣地區鄉鎮市區分為：工商市鎮、新興鄉鎮、綜合性市鎮、服務性鄉鎮、坡地鄉鎮、偏遠鄉鎮、山地鄉鎮，我們剔除分類中屬外島地區的鄉鎮，將偏遠鄉鎮及山地鄉鎮合併為一層，並加上台北市、高雄市、省轄市共九層。

N 為母體總數， n 為樣本總數， N_h 為各分層母體總數， n_h 為各分層樣本數， α_h 為第 h 分層抽取的總鄉鎮數， b_h 為第 h 分層各鄉鎮抽取的村里數， c_h 為第 h 分層各村里的總抽取人數。
 α 為第 α 個鄉鎮， $\beta_{h\alpha}$ 為第 h 分層第 α 鄉鎮的總人口數， $c_{h\alpha\beta}$ 為第 h 分層第 α 鄉鎮第 β 村里的總人口數。

表二：抽樣母體資訊

分層	母體人數	鄉鎮數	抽出鄉鎮數	抽出樣本數	使用樣本數	完訪樣本數
新興鄉鎮	627,687	43	8	902	571	279
工商市鎮	573,823	9	5	768	523	254
綜合性市鎮	412,950	22	6	594	394	198
坡地鄉鎮	240,865	54	3	354	237	115
山地及偏遠鄉鎮	530,688	121	9	846	381	269
服務性鄉鎮	431,876	51	6	623	363	213
省轄市	485,752	26	6	699	389	232
台北市	549,850	12	5	921	672	263
高雄市	290,261	11	2	474	223	135
總計	4,143,752	349	50	6,181	3,753	1,958

2. 反覆多重加權 (Raking)

在完成不等機率加權後，為使成功樣本結構具有代表性並符合母體結構，隨即針對「性別」、「年齡」(七分類)、「教育程度」(五分類)與「地區層別」(九分類)四個變項進行樣本代表性檢定(卡方檢定)，並採用「多變數反覆加權法」進行加權，直到樣本代表性檢定結果符合母體的分佈狀況為止。

最終加權權值的計算公式：
$$w_{final} = w * \frac{N_{性別}}{n_{性別}} * \frac{N_{年齡}}{n_{年齡}} * \frac{N_{地區層別}}{n_{地區層別}} * \frac{N_{教育程度}}{n_{教育程度}} * \frac{n}{N}$$

由表三，本研究之成功樣本代表性檢定顯示，成功樣本在「性別」、「年齡」及「教育程度」上的結構明顯與母體不一致。為了與母體結構更符合，避免造成分析資料時推論的偏差，故以「多變數反覆加權法(raking)」對成功樣本進行加權。

由表四，經加權處理後，成功樣本在「性別」、「年齡」、「教育程度」及「地區層別」的分佈上，均與母體分佈無差異。

表三：成功樣本代表性檢定(加權前)

		樣本		母體	檢定結果	
		人數	百分比	百分比	卡方值	p 值
性別	男	925	47.24	49.93	5.664	0.017
	女	1,033	52.76	50.07		
年齡	46-48 歲	399	20.38	23.10	104.048	0.001
	49-51 歲	321	16.39	17.62		
	52-54 歲	198	10.11	14.88		
	55-57 歲	304	15.53	11.98		
	58-60 歲	318	16.24	11.61		
	61-63 歲	249	12.72	10.87		
	64-66 歲	169	8.63	9.93		
地區層別	新興鄉鎮	279	14.25	15.15	3.999	0.857
	工商市鎮	254	12.97	13.85		
	綜合性市鎮	198	10.11	9.97		
	坡地鄉鎮	115	5.87	5.81		
	山地及偏遠鄉鎮	269	13.74	12.81		
	服務性鄉鎮	213	10.88	10.42		
	省轄市	232	11.85	11.72		
	台北市	263	13.43	13.27		
	高雄市	135	6.89	7.00		
教育程度	小學及以下	1,259	64.30	57.76	35.921	0.001
	初(國)中/初職	221	11.29	12.82		
	高中職	277	14.15	16.41		
	專科	82	4.19	5.60		
	大學及以上	119	6.08	7.42		

樣本總數：1,958

表四：成功樣本代表性檢定(加權後)

		樣本		母體	檢定結果
		人數	百分比	百分比	
性別	男	978	49.93	49.93	$\chi^2=0.00196305$ $p > 0.05$ 樣本與母群一致
	女	980	50.07	50.07	
年齡	46-48 歲	453	23.14	23.10	$\chi^2=0.00748121$ $p > 0.05$ 樣本與母群一致
	49-51 歲	345	17.64	17.62	
	52-54 歲	291	14.87	14.88	
	55-57 歲	234	11.97	11.98	
	58-60 歲	227	11.61	11.61	
	61-63 歲	213	10.85	10.87	
	64-66 歲	194	9.92	9.93	
地區層別	新興鄉鎮	296	15.14	15.15	$\chi^2=0.00525898$ $p > 0.05$ 樣本與母群一致
	工商市鎮	271	13.85	13.85	
	綜合性市鎮	195	9.96	9.97	
	坡地鄉鎮	114	5.80	5.81	
	山地及偏遠鄉鎮	250	12.79	12.81	
	服務性鄉鎮	204	10.41	10.42	
	省轄市	230	11.74	11.72	
	台北市	261	13.31	13.27	
	高雄市	137	7.00	7.00	
教育程度	小學及以下	1,131	57.76	57.76	$\chi^2=0.00164159$ $p > 0.05$ 樣本與母群一致
	初(國)中/初職	251	12.82	12.82	
	高中職	321	16.41	16.41	
	專科	110	5.60	5.60	
	大學及以上	145	7.42	7.42	

樣本總數：1,958

權數使用說明：

Stata（測試版本為 13）：

使用前須打入以下設定語法：

```
svyset PSU [pw=weight_adj_3], strata(strata)
```

weight_adj_3 即為本單位所提供之最終權數，PSU（即此樣本的第一抽樣單位，Primary Sampling Unit (PSU)）為此樣本的 PSU 變數，strata 為此樣本的分層變數。完成設定後，在使用語法前，須在前方加上「svy:」。以性別變項（a01）為例，若要檢查其次數分配，須使用以下語法：svy: ta a01

SAS（測試版本為 9.2）：

使用時須打入以下設定語法（以性別變項 a1 為例）：

```
proc surveyfreq data=data2016;
```

```
tables a01;
```

```
strata strata;
```

```
cluster PSU;
```

```
weight weight_adj_3;
```

```
run;
```

strata 處放置樣本的分層變數，cluster 處放置 PSU 變數，weight_adj_3 即為本單位所提供之最終權數。

R：

事前須安裝以下套件

```
library(survey)
```

使用前亦須打入以下設定語法：

```
svydata1999 <- svydesign(id=~PSU, weights=~ weight_adj_3, strata=~strata,  
data=data1999, nest=TRUE)
```

PSU 為此樣本的 PSU 變數；weight_adj_3 為本單位所提供之最終權數；strata 為此樣本的分層變數；data 為原資料檔名稱。若要進行描述統計（以性別變項為例），可使用以下語法：

```
prop.table(svytable(~a01, design = svydata1999))
```

SPSS（須包含「複式樣本」功能之版本方可執行）：

設定步驟：分析 → 複式樣本 → 準備分析 → 選取「建立計畫檔案」後瀏覽存取檔案，完成後按下一步 → 點選 strata 變項至「階層」欄位 → 點選 PSU 變項至「集群」欄位 → 點選 weight_adj_3 權數變項至「樣本加權」欄位，完成

後按下一步 ➡ 選取「WR（放回取樣）」後按下一步 ➡ 接著可依個人需求選擇「將規格儲存至計劃檔案」或「將精靈產生的語法貼到語法視窗中」，點選後按「完成」即可。

完成上述設定步驟後，可進行分析，步驟如下（以性別變項為例）：

分析 ➡ 複式樣本 ➡ 次數分配表 ➡ 於「計畫-檔案」中進行瀏覽，選取剛剛設定時所建立的計畫檔案，選取後按「繼續」 ➡ 點選 a01 至「次數表格」中 ➡ 點選右側的「統計資料」後，選取「表格百分比」與其他個人所需之統計資料，選取完成後點選「繼續」 ➡ 接著點選左下方之「確定」，即可於輸出視窗獲得相關統計資訊。