

100 年婦女生活狀況調查抽樣設計及母體推估

一、抽樣規劃

(一) 樣本數

本調查訪問 6,455 份有效樣本，在信心水準為 95% 的情況下，抽樣誤差在 ± 1.22 個百分點之內。各縣市最低樣本數為 138 份，最大樣本數為 1,039 份，在 95% 的信心水準下，各縣市推估之抽樣誤差界於 $\pm 8.34\%$ 至 $\pm 3.04\%$ 之間。

(二) 抽樣母體

以電話住宅用戶為抽樣母體。

以臺閩地區各鄉鎮市區電話區局碼為抽樣母體，採用電話號碼後 4 碼隨機方式抽樣，即使電話戶並未將電話號碼登記，仍保有相等機率被抽中並進行訪問。

訪問過程中會同時接觸非住宅用戶電話，則將在問卷開始前即過濾，確認該筆電話為住宅用戶(住商合一)電話才會繼續進行後續戶中抽樣及訪問工作。

(三) 抽樣方法

1. 分層準則與層數：先將臺閩地區依照統計區域分為下列 5 層。

(1) 北部地區：新北市、臺北市、基隆市、新竹市、宜蘭縣、

桃園縣、新竹縣

(2) 中部地區：臺中市、苗栗縣、彰化縣、南投縣、雲林縣

(3) 南部地區：高雄市、嘉義市、臺南市、嘉義縣、屏東縣、

澎湖縣

(4)東部地區：花蓮縣、臺東縣

(5)金馬地區：金門縣、連江縣

2.抽樣方法：採分層比例隨機抽樣法；每層之各縣市依照縣市內 15 至 64 歲女性人口數占臺閩地區 15 至 64 歲女性總人口數(含調查期間在臺而尚未取得我國國籍之外籍配偶及大陸、港澳籍配偶)的比例分配樣本數。

(四) 樣本配置

樣本配置方面採比例配置法，依據內政部 100 年 6 月公布我國各縣市 15-64 歲女性人口數比例決定各層應抽樣本數，力求分析穩定性並在有限預算下，若配置後若縣市樣本數未達 138 份則增補至 138 份（金門縣及連江縣共配置 138 份），可使各縣市最大抽樣誤差在 8.34% 以內，先以總樣本數 5,385 份進行配置，增補完總樣本數為 6,200 份，樣本配置之計算方法如下：

$$n_i = \frac{N_i}{N} \times 5,385$$

N：15-64 歲女性人口總數（民國 100 年 6 月）

N_i ：某地區層 15-64 歲女性人口數（民國 100 年 6 月）

n_i ：第 i 地區層應抽樣本數

i 為地區層，i=1 為新北市、i=2 為臺北市、...、i=22 為連江縣

表 1-1 分層樣本結構

民國 100 年 6 月

地區別	15-64 歲 人口數	15-64 歲 人口比例	依比例 樣本配置	增補後 樣本數	樣本配置 抽樣誤差
總計	8,670,326	100.00	5,385	6,200	1.24
新北市	1,545,826	17.83	960	960	3.16
臺北市	1,018,190	11.74	632	632	3.90
臺中市	1,011,117	11.66	628	628	3.91
臺南市	694,559	8.01	431	431	4.72
高雄市	1,057,578	12.20	657	657	3.82
宜蘭縣	163,536	1.89	102	138	8.34
桃園縣	760,260	8.77	472	472	4.51
新竹縣	180,290	2.08	112	138	8.34
苗栗縣	194,080	2.24	121	138	8.34
彰化縣	458,188	5.28	285	285	5.81
南投縣	183,320	2.11	114	138	8.34
雲林縣	237,018	2.73	147	147	8.08
嘉義縣	181,132	2.09	112	138	8.34
屏東縣	311,836	3.60	194	194	7.04
臺東縣	78,645	0.91	49	138	8.34
花蓮縣	120,532	1.39	75	138	8.34
澎湖縣	33,965	0.39	21	138	8.34
基隆市	144,391	1.67	90	138	8.34
新竹市	154,278	1.78	96	138	8.34
嘉義市	101,020	1.17	63	138	8.34
金門縣	37,363	0.43	23	138	8.34
連江縣	3,201	0.04	2		

二、樣本代表性檢定

本次調查實際完成 6,455 份有效樣本，在信心水準為 95% 的情況下，抽樣誤差在 ± 1.22 個百分點之內。本次調查過程在地區上有增補樣本，為求樣本結構與母體結構相符，因此除依據母體地區結構進行樣本結構加權調整外，並同時針對樣本之年齡及教育程度進行加權調整。加權前後之樣本代表性檢定詳見表 1-2 至表 1-4。

加權方式採多變項逐步加權(Raking)法，加權過程如下：

步驟一：首先就各地區別下之教育程度及年齡層分別計算出母體比例，並以此比例調整樣本結構。第一步先調整樣本之教育程度，第二步在第一步之權數下調整樣本之年齡層。

步驟二：反覆上述三次兩步驟的調整後檢定樣本地區別、年齡及教育程度分配與母體分配均已無顯著差異，故停止 raking 調整步驟。每一筆資料都是乘以調整權數，使調整後樣本權數與母體的分配完全一致。最後的權數是各步調整權數累乘。權數(W)計算如下：

$$W = \frac{N_i}{N} \bigg/ \frac{n'_i}{n}$$

N_i 和 n'_i 是第 i 交叉組的母體人數和樣本加權人數

N 和 n 是母體總人數和樣本加權總人數

表 1-2 樣本地區別與母體年齡結構一致性檢定

民國 100 年 6 月

單位：人數；%

地區別	母體結構		加權前樣本分配		加權後樣本分配		卡方檢定
	人口數	百分比	樣本數	百分比	樣本數	百分比	
總計	8,670,326	100.00	6,455	100.00	6,455	100.00	加權前：
新北市	1,545,826	17.85	1,039	16.10	1,152	17.85	卡方值為 1155.72 > 31.41 (自由度為 20, 顯著度為 5%)
臺北市	1,018,190	11.78	683	10.58	760	11.78	
臺中市	1,011,117	11.70	683	10.58	755	11.70	在 5%顯著水準下，樣本與母體的地區別結構有顯著差異。
臺南市	694,559	8.04	466	7.22	519	8.04	
高雄市	1,057,578	12.21	712	11.03	788	12.21	加權後：
宜蘭縣	163,536	1.89	145	2.25	122	1.89	
桃園縣	760,260	8.72	511	7.92	563	8.72	卡方值為 0.00 < 31.41
新竹縣	180,290	2.06	139	2.15	133	2.06	
苗栗縣	194,080	2.23	138	2.14	144	2.23	在 5%顯著水準下，樣本與母體的地區別結構無顯著差異。
彰化縣	458,188	5.28	310	4.80	341	5.28	
南投縣	183,320	2.11	140	2.17	136	2.11	
雲林縣	237,018	2.72	158	2.45	175	2.72	
嘉義縣	181,132	2.07	139	2.15	134	2.07	
屏東縣	311,836	3.58	213	3.30	231	3.58	
臺東縣	78,645	0.91	139	2.15	58	0.91	
花蓮縣	120,532	1.39	142	2.20	90	1.39	
澎湖縣	33,965	0.39	138	2.14	25	0.39	
基隆市	144,391	1.66	142	2.20	107	1.66	
新竹市	154,278	1.78	139	2.15	115	1.78	
嘉義市	101,020	1.17	141	2.18	75	1.17	
金馬地區	40,564	0.47	138	2.14	31	0.47	

資料來源：母體結構人口數為內政部戶政司統計本國籍 8,535,167 人及內政部入出國及移民署統計調查期間推計在臺而尚未取得我國國籍之外籍配偶 67,767 人及大陸、港澳籍配偶 67,392 人之 15 至 64 歲女性合計。

表 1-3 樣本年齡別與母體年齡結構一致性檢定

民國 100 年 6 月

單位：人數；%

年齡別	母體結構		加權前樣本分配		加權後樣本分配		卡方檢定
	人口數	百分比	樣本數	百分比	樣本數	百分比	
總計	8,670,326	100.00	6,455	100.00	6,455	100.00	加權前：
15-19 歲	786,063	9.07	585	9.07	585	9.07	卡方值為 4.45<16.92 (自由度為 9，顯著 度為 5%) 在 5% 顯著 水準下，樣本與母體 的年齡別結構無顯 著差異。
20-24 歲	778,735	8.98	580	8.98	580	8.98	
25-29 歲	924,311	10.66	688	10.66	688	10.66	
30-34 歲	1,037,171	11.96	772	11.96	772	11.96	
35-39 歲	923,183	10.65	687	10.65	687	10.65	
40-44 歲	940,944	10.85	701	10.85	701	10.85	加權後： 卡方值為 0.00<16.92 在 5% 顯著水準下， 樣本與母體的年齡 別結構無顯著差異。
45-49 歲	967,499	11.16	720	11.16	720	11.16	
50-54 歲	914,544	10.55	681	10.55	681	10.55	
55-59 歲	820,984	9.47	611	9.47	611	9.47	
60-64 歲	576,893	6.65	429	6.65	429	6.65	

資料來源：內政部戶政司，「各縣市人口數按性別及單一年齡分」(100 年 6 月)、內政部入出國及移民署提供調查期間在臺而尚未取得我國國籍之外籍配偶及大陸、港澳籍配偶(100 年 6 月)。

表 1-4 樣本教育程度別與母體教育程度結構一致性檢定

民國 100 年 6 月

單位：人數；%

教育程度別	母體結構		加權前樣本分配		加權後樣本分配		卡方檢定
	人口數	百分比	樣本數	百分比	樣本數	百分比	
總計	8,494,504	100.00	6,455	100.00	6,455	100.00	加權前：
小學以下	987,445	11.62	484	7.50	737	11.42	卡方值 266.36 > 11.07 (自由度為 5， 顯著度為 5%) 在 5% 顯著水準下， 樣本與母體的教育 程度結構有顯著差 異。
國初中	1,153,059	13.57	683	10.58	872	13.51	
高中職	2,880,749	33.91	2,182	33.80	2,182	33.81	
專科	1,171,145	13.79	861	13.34	888	13.76	
大學	1,963,647	23.12	1,917	29.70	1,491	23.10	
研究所	338,459	3.98	297	4.60	257	3.98	加權後： 卡方值為 0.16<11.07 在 5% 顯著水準下， 樣本與母體的教育 程度結構無顯著差 異。
不知道	-	-	31	0.48	27	0.42	

資料來源：內政部統計處，「各縣市十五歲以上現住人口數按性別、年齡及教育程度分」(99 年)

三、資料處理及分析

(一) 資料處理

1.電話訪問的複查：針對邏輯有誤的樣本進行複查，若情況嚴重，則另行訪問，重新補回所需樣本。

2.電話訪問的資料檢誤

原始資料經過複查後，依據事先列出的檢誤表，透過 SPSS 軟體撰寫程式，找出有缺失的資料，以電腦進行資料檢誤的動作，確保每一筆資料都符合邏輯性。若某筆資料經檢誤發現邏輯性錯誤時，則實際回撥給受訪者確認實際答案。

3.遺漏值的處理

遺漏值發生時，除了「不知道/無意見」之答案外（意見性問項常有此種無法作答的情況），若單筆資料有二分之一以上問項未回答，或幾個特別重要的問項未回答，則回撥受訪者補齊答案，或直接刪除該樣本找尋條件相似（如：相同居住地區、年齡、性別）的替代樣本重新訪問。若單筆資料經重新撥回詢問，但受訪者仍拒答或無法與受訪者接觸，則刪除該份樣本，已遺漏值的情形處理。此次調查有 1,944 筆二分之一以上問項未回答，主要為受訪者中途拒答，因此刪除該樣本找尋條件相似的替代樣本重新訪問；有 66 筆資料少部分問項未回答，皆經重新撥回詢問，占整體比例約 1.02%。

(二) 統計分析方法

不同狀況下使用的分析方法如下：

1.次數分配

根據各題加權的樣本比例進行比較選項間的差異，用下列 Z-檢定，檢驗兩選項間百分比(P_1 和 P_2)的差異：

$$Z_1 = \frac{P_1 - P_2}{\sqrt{\frac{1}{n}[P_1 + P_2 - (P_1 - P_2)^2]}}$$

2.交叉分析

以各題與基本資料的交叉表來分析民眾對各議題的看法與他們基本特徵間的相關。交叉表第一步採用卡方檢定，交叉表的卡方顯著水準小於 5%時才認定兩變數間有相關。

$$\chi^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad \begin{cases} O_{ij} \text{ 為第 } i \text{ 列第 } j \text{ 行觀察值} \\ E_{ij} \text{ 為第 } i \text{ 列第 } j \text{ 行期望值} \end{cases}$$

四、母體推估

調查結束後，進行樣本結構與母體結構適合度檢定(goodness of fit test)，由於部分縣市在抽樣時已增加樣本，因此與母體結構不符，故採用地區、年齡及教育程度加權方法調整以期與母體結構相符。經過權數調整後，母體推估方式如下：

(一) 母體整體百分比的推估

$$P_{st} = \frac{\sum_{j=1}^J \sum_{k=1}^{n_j} y_{stjk} w_{1jk}}{n}$$

其中， P_{st} = 第 s 題第 t 項特徵的比例

$$y_{stjk} = \begin{cases} 1, \text{第 } j \text{ 地區層第 } k \text{ 個樣本在第 } s \text{ 題具有 } t \text{ 項特徵} \\ 0, \text{第 } j \text{ 地區層第 } k \text{ 個樣本在第 } s \text{ 題不具有 } t \text{ 項特徵} \end{cases}$$

w_{1jk} = 第 j 地區層第 k 個樣本的權數(考慮地區別、年齡、教育程度結構)

n_j = 第 j 地區層內有效樣本數

n = 有效樣本總數

j 為地區層，j=1 為新北市、j=2 為臺北市、...、j=22 為連江縣

(二) 母體總數的推估

母體總數(X)點估計

$$\hat{X} = N \times \bar{\mu}$$

式中 \hat{X} ：母體總數點估計值

N：母體總人數

$\bar{\mu}$ ：母體總平均數點估計值