



## 一、抽選樣本戶：母群清冊型式考量

在「訪問方法-母群清冊型式」考量方面，係指不同的訪問方法、會有不同的母群清冊，而發展出不同的抽選樣本戶的方法。

訪問的方法有4類：「電訪、造訪、函訪、網訪」。造訪的清冊是戶籍資料或地圖；函訪是戶籍資料；而網訪則無法取得公眾的實體清冊，也不易設計虛擬清冊。

當前調查實務仍以電話訪問為多，故以這個方法為核心討論。

## 二、電話調查抽樣法

在市內電話-即有線電話未百分之百普及前，有些人不能被電話連繫到，此清冊即不完整，便違反了隨機/等機率性的假設。

美國的有線電話普及率於 1975年達到90%，臺灣則於1980年後達到90%，在其後幾年甚至接近百分之百普及率。

因此，造成當時是電話調查準確度的黃金時期，本研究者也是在此時期[在臺灣導入電話調查](#)。

唯2000年後陸續發生的2大社會行為趨勢，會使研究者對電話號碼掌握不全，有些樣本永不會被抽中，等同母群清冊不完整。

這2大社會行為趨勢是：電話簿不登記者增加，以及行動電話的衝擊。

電話調查抽樣過去主要有三類方法：

1980~電話調查初期：電話簿法。

1990~電話簿不登記者增加時期：隨機撥號法（Random Digit Dialing,簡稱 RDD）以及綜合法。

而2010年代行動電話興起，市內電話調查效益大幅降低，可能的解決方案有：

當前的網路大數據分析補強法。

展望未來的：「行動調查法」：使用「Google 地球」，有可能形成史上最有效益的清冊。

數位電視機上盒法：在方法論的基礎上革新，根本取消「訪問程序」，而發展「數位行為調查法」。

### (一)電話簿法

以電話簿作為實體的母群清冊。



## 1、電話簿系統抽樣

過去最主要的電話簿抽樣方法就是「系統抽樣」(吳統雄 [35])。

系統抽樣的程序是(範例5-1)：

a. 根據電話簿的頁數及樣本大小，決定平均每間隔若干頁抽出一個樣本，假設這個間隔為  $n$ 。

b. 在一至  $n$  的數字中，隨機抽出一個數字，然後以這個數字為第一頁，每隔  $n$  頁抽出樣本戶所在的頁數。

c. 隨機抽出第幾欄、第  $m$  個號碼為樣本戶，每一頁均以這個位置的號碼為準。

d. 如果電話簿中每一欄中電話號碼的數量相差很大時--譬如調查家具業喜愛用那一種木料，必需借助分類電話號碼簿，但是分類簿中常常插有廣告，使得  $m$  可能大於某欄的樣本數，則跳到下一欄，繼續數到  $m$  的位置。

也有學者建議，每一欄所有的號碼不一定時，那麼就可以事先決定：有一半的樣本是由每一欄的上方往下數若干個號碼後選出，另外一半的樣本則是由每欄的下方往上數而選出。不過根據統計學的觀點，由於插放廣告也是一種純粹隨機/等機率的行為，Madow et al. [238] 以數理證明這樣的顧慮也可以是不必要的。

### 範例5-1 電話簿抽樣法

#### (問題)

某研究者欲在臺北市用電話簿抽出 400個電訪樣本，應如何抽選？

#### (已知)

- 1、電話簿計1762頁。
- 2、每頁有 4 欄號碼。
- 3、每欄號碼約有 85至87個號碼。

#### (方法)

- 1、以1762 除400，得 4.4，推定每 4 頁抽一個號碼。
- 2、在由 1 至 4 之間，隨機抽出一個號碼，假設為 3，則以 3 為第一頁，每隔 4 頁抽出一頁，即抽 3，7，11，15……等，計400 頁。
- 3、在由 1 至 4 之間，隨機抽出一個號碼，假設為 2，則每頁之第 2 欄，為樣本所在之欄。
- 4、每欄號碼最多為 87，故在 1 與 87 之間隨機抽出一個號碼，假設為 54，則每欄之第 54 個號碼便為樣本。計得 400 個樣本。如果抽到 86、或 87，但該頁只有 85 或 86 個號碼，則以該頁號碼數隨機「更新」重抽。「更新」是指母群清冊中沒有這個樣本，所以重抽不會影響其他樣本被抽中的機會。
- 5、以 440 除 400，得 1.1；推定每 11 頁要剔除一個號碼，仍以上述隨機方式剔除之，得所需之 400 個樣本。剩餘 40 個樣本，若發生不存在樣本(如記載錯誤、遷移等)時，可作為更新樣本使用。



## 2、電話簿簡單隨機抽樣

Blankenship [103]曾經推介在使用電話簿時，採用簡單隨機抽樣法，就是藉亂數表在電話簿中抽出樣本；因為電話簿的排列並沒有系統性的偏差，在電腦尚未普及之前，為了節省工作負擔起見，調查一般均使用系統抽樣。而本項資訊系統則可輕易提供簡單隨機抽樣的選擇。

應用電話簿法的優點：容易實施。

應用警惕：抽樣代表性與電話普及率成正比，如果電話普及率很低，將無隨機性。

Blankenship [103]也指出，如果根據電話簿抽樣，有三種特殊的情形，也可能損及電話訪問的代表性：

- (1) 電話碼記載不正確；
- (2) 有的人電話「不登記」；
- (3) 有的電話號碼也許會變更。

另外，還要顧慮電話簿中可能沒有涵蓋全體電話戶、以及新近安裝話機的人。

### (二)隨機撥號法

基於這種考慮，許多學者發展出了「電話隨機撥號」（Random Digital Dialing, RDD）技術，嘗試克服這個困難。[136][351]

這個方法就是把所有電話號碼的組合，視為一個「虛擬母群清冊」，那麼每個號碼都有被抽中的機會。

「隨機撥號」技術，其中又可分為簡單隨機撥號法（Simple RDD）和集群隨機撥號法（Cluster RDD）。

#### 1、簡單隨機撥號

以臺北市的電話為例，每個號碼一共有8碼，前4個是分區號，後四個則是戶號。隨機撥號的程序是：

- a.先選定欲訪問的分區號碼。
- b.以隨機/等機率的方式，選定後四碼戶號。

#### 2、集群隨機撥號

隨機撥號發展以後，發現一個流弊，就是電話的「空號」太多，減低了電話訪問的效率，又增加了開支，與電訪的最佳效益正好大相逕庭，於是又有了集群隨機撥號的構想，這個方法的程序是：



- a. 每一百個號碼劃為一群，因此每一個區域號下便有一百群。
  - b. 以比例抽樣方法決定每一個區域號碼中可以抽出若干個樣本群。
  - c. 以隨機抽樣抽出若干個群，以隨機撥號來試打，如果某一群的第一通電話不是空戶，這個群就保留為樣本群，再抽出群內的樣本；如果第一通電話是空戶，就放棄整個群，另外抽一群來代替。
- 這樣處理的方式，Cumplings [143]發現有效樣本增加三分之一，Groves [186]發現增加三倍，而資料的變異數與簡單隨機法接近相同。

### (三)綜合法

綜合法是要兼取電話簿法和隨機撥號法兩者之長，過去有「隨機雙數」( 2 Random Digits )法和「加一法」( plus-one )兩種，本研究則再建議改良為「隨機尾數」法。

#### 1、隨機雙數

Hauck et al. [189]認為，如果先從電話簿上抽出一個號碼，然後隨機/等機率改變最後的兩位數字，那麼發生空號的情形就會比較少，而且也達到隨機撥號的目的，經過他們實際的操作，效果也不錯。

#### 2、加一法

「隨機雙數」的構想很好，可是在人工時代，每抽出一個號碼後，還要再抽出兩個隨機數，手續上還是麻煩了一點。Landon et al. [230]有鑑於此，便建議每當在電話簿中抽出一個號碼後，再把它加一才變成樣本戶，譬如抽出的號碼是 27641234，訪問號碼卻是 27641235，也可以達到隨機撥號的目的。這個方法經過吳統雄[35]於1980年代在國內的實驗發現未接通的平均比率是 31.60%，比較系統抽樣的未接通率 22.19%，加一法約高出了三成。

加一法在1985年的完成狀況由1983年的72.2%降為64.6%，而聯繫不到的由23.3%升到28.1%，可能是1985年抽樣地區多在台灣省，而台灣省由加一法遇到的空戶較多。不過，加一法的損失不能算太嚴重，在經費充足的時侯，這樣的損失尚可忍受，效果堪稱良好。

加一法的優點是易於手工實施，但是，如果電話不登記的人，遠遠超過登記的人時，有可能連號的人都沒有登記，那麼還是會發生不隨機/等機率的情形。

### (四)電話簿法和隨機撥號法的比較與取捨



集群隨機撥號雖可以改善簡單隨機撥號的效益，但是比起用電話簿抽樣，還是不經濟了很多。Rich [286]早期曾經進行大規模調查，發現當時用電話簿抽出來的樣本不論在人口特徵上，或是對各方面的意見主張上，已經很能夠代表全體人口了，用隨機撥號所能增加的統計精密度很有限，但卻所費不貲，因此不贊成使用隨機撥號。

Sudman [338]也有相同的發現，他的建議則是：當母群體很大，而不在電話簿上的人口在比例上很少時，使用電話簿抽樣即可，反之，則應使用隨機撥號。

雖然早期有學者認為，電話簿所提供的清冊內容已經大得有足夠的代表性，可以泯滅「不登記」等情形帶來的偏差。

但什麼叫做「足夠大」？文獻上並無可信服的實驗證據，僅有直觀的，以完整性達到90%，就是「足夠大」。另有多項文獻指出，在電話簿上「登記」和「不登記」的人，人口與行為特徵並不相同。

本研究者的研究經驗顯示，臺灣自1996年電信自由化，電信局改制中華電信後，「不登記」的人大幅增加，亦即隨機/等機率性大幅下降，使用隨機撥號法更有其必要。

以[2012總統選舉預測](#)為例，5家民調機構均預測蔡英文勝，平均數字為蔡勝 9.68%，而吳統雄預測馬勝 7%，中選會公布票數，實際馬勝 6.57%。

亦即所有坊間民調機構的單側誤差達到17%，而吳統雄的單側誤差僅為：0.4%。其中重要原因之一，就是吳統雄採用了嚴謹的「資訊系統隨機尾數撥號法」。

## (五)資訊系統隨機尾數法

在資訊系統協助下，產生隨機數便不困難。

臺灣地區電話不登記比例，估計在都會已超過50%，在鄉鎮地區也達到33%-全體三分之一以上，亦即母群清冊不完整的情形相當顯著，若採用一般的電話簿抽樣法，將嚴重傷害抽樣隨機/等機率性與樣本的代表性。因此採用可涵蓋所有樣本的隨機撥號法，應有相當必要。

在經過實證研究後，在各種隨機撥號法中，建議採用本研究者所發展的資訊系統隨機尾數法，而以單尾數或雙尾數為優先適用。

## (六)行動電話興起的挑戰與回應

2010年代行動電話興起，市內電話調查效益大幅降低，可能的解決方案有：

當前的網路大數據分析補強法。

展望未來的：「行動調查法」：使用「Google 地球」，有可能形成史上最有效益的清冊。



數位電視機上盒法：在方法論的基礎上革新，根本取消「訪問程序」，而發展「數位行為調查法」。

美國疾病控制與預防中心([Centers for Disease Control and Prevention's, CDC](#))與美國國家衛生統計中心([National Center for Health Statistics, NCHS](#))有一個年度合作研究計畫：國家健康調查([National Health Interview Survey, NHIS](#))，過去數十年都是採用市內電話調查，在行動電話興起後，為了調整研究方法的目的，發展了一項子計畫：行動電話取代行為調查(**Wireless Substitution**)，足以反映行動電話興起的挑戰規模，以及啟示研究者思考如何回應。

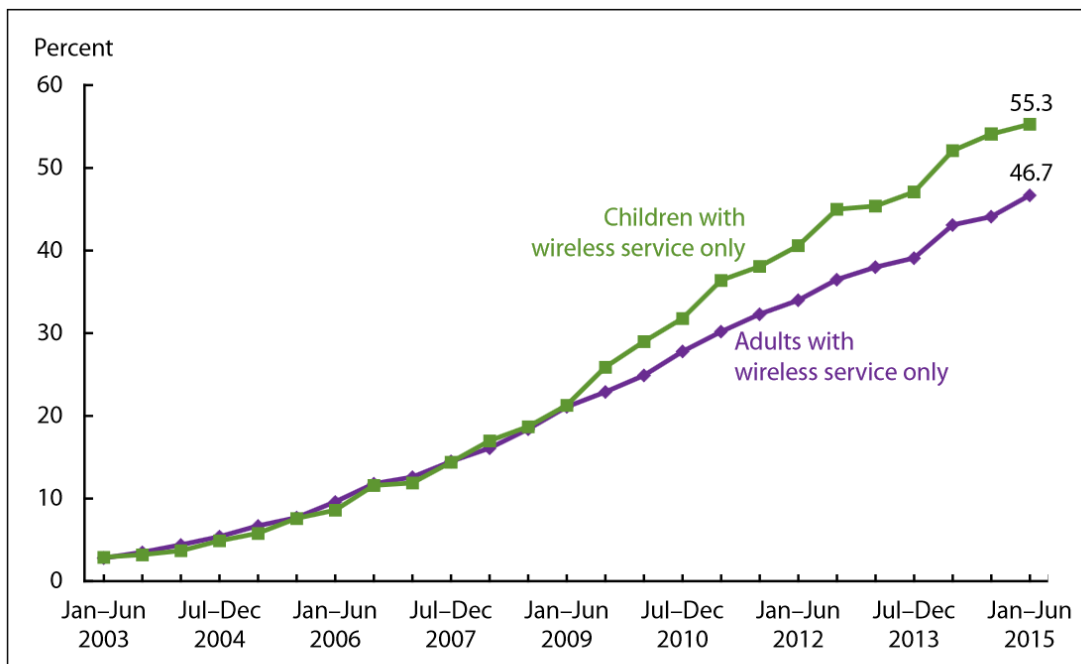
臺灣似乎並無類似研究，故有他山之石，可以攻錯的參考價值。

## 1、行動電話超過市內電話

根據最新(2015)的行動電話取代行為調查([Wireless Substitution, 2015](#))報告顯示，美國家庭中 47.4% 只使用行動電話，根本就沒有安裝有線市內電話。

在2014年上期，美國家庭使用行動電話與有線市內電話者，百分比相當，而至2014下期，使用行動電話者首度超過使用有線市內電話者，比率為：45.4：42.7。而2015變成：47.4：41.6。

就使用者分，成人（18歲以上）只用行動電話的，為所有成人的46.7%，約1億1千3百萬人；而兒童更高達所有兒童的55.3%，約4千1百萬人。目前美國住家擁有市內有線電話的比率，只和20世紀中、約第二次世界大戰的時期差不多。



NOTE: Adults are aged 18 and over; children are under age 18.  
DATA SOURCE: CDC/NCHS, National Health Interview Survey.



## 2、行動電話使用者的人口特徵

Demographic characteristic	January–June 2015	95% confidence interval <sup>1</sup>
<b>Race/ethnicity</b>		
Hispanic or Latino, any race(s)	59.2	56.98–61.33
Non-Hispanic white, single race	43.2	41.57–44.80
Non-Hispanic black, single race	48.1	45.70–50.51
Non-Hispanic Asian, single race	47.9	44.01–51.77
Non-Hispanic other, single race	51.8	43.35–60.14
Non-Hispanic multiple race	53.6	46.53–60.54
<b>Age (years)</b>		
18–24	59.4	56.95–61.90
25–29	71.3	68.92–73.52
30–34	67.8	65.31–70.14
35–44	56.6	54.85–58.32
45–64	40.8	39.24–42.40
65 and over	19.3	17.92–20.75
<b>Sex</b>		
Male	48.2	46.92–49.56
Female	45.3	43.92–46.62
<b>Education</b>		
Some high school or less	49.0	46.87–51.22
High school graduate or GED <sup>2</sup>	46.7	44.93–48.41
Some post-high school, no degree	49.0	47.15–50.87
4-year college degree or higher	43.5	41.55–45.45
<b>Employment status last week</b>		
Working at a job or business	52.7	51.26–54.15
Keeping house	49.3	46.54–52.12
Going to school	49.6	45.91–53.33
Something else (incl. unemployed)	32.7	31.21–34.16
<b>Household structure</b>		
Adult living alone	51.1	49.02–53.15
Unrelated adults, no children	84.6	78.68–89.13
Related adults, no children	39.1	37.43–40.79
Adult(s) with children	53.3	51.62–55.06
<b>Household poverty status<sup>3</sup></b>		
Poor	59.3	56.44–62.03
Near-poor	54.4	51.92–56.93
Not-poor	45.7	44.03–47.29

這份報告中最值得注意的是行動電話使用者的人口特徵。哪些人使用有線電話，而哪些人不使用有線電話？



根據這份調查報告，以使用行動電話為主、亦即在分類中超過505的族群共有五大類：

(1) 單身者

沒有家屬同住、也沒有子女的成年人，84.6%。

(2) 年輕人

44歲以下的美國人，愈年輕愈多，56.6~71.3%。

(3) 窮人

家境清寒的成年人，59.3%。

(4) 拉丁族裔

拉丁族裔是各種族最多者，59.2%

(5) 就業者

在正常上班、或自行執業者，52.7%。

另外，居住條件不穩定就不會裝有線電話、優渥家庭則兩者皆備。居住條件越穩定，才越可能使用有線電話。若是時常搬家，使用有線電話將十分麻煩，此數據也反映了美國當前的居住生態。

家境優渥的美國家庭，較多同時使用行動電話與市內電話，兩者可以兼得。

### 3、行動電話使用者的健康特徵

由於這是國家健康調查(National Health Interview Survey, NHIS)研究，所以也有行動電話使用者的健康特徵報告，而發現了幾項特殊現象：

行動電話使用者與沒有電話者，比使用市內電話者，較有飲酒習慣。

使用市內電話者，比行動電話使用者與沒有電話者，較多具有防疫、有病就醫的行為。

從以上的比較，可知使用市內電話者，與行動電話使用者，應有其他不同取用行為的可能。

當前的各種抽樣調查，如果無法接觸只使用行動電話的使用者，無法排除產生重大誤差。

針對以上挑戰，可能有以下的解決方案。



### (七)網路大數據分析補強法

在各種網路大數據分析法中，Google 排行榜提供了一個極佳的「關注程度」母群清冊，配合內容分析法，可以觀察出相當多現象。這種方法無法具體測量樣本個人行為，但等於是對一個群體行為的測量，在分析民意上，有補強的作用與效果。

### (八)行動調查法

社會調查史上，最早使用的母群清冊是地圖，而「Google 地球」提供了數位網路時代的即時地圖。

展望未來如何提升「Google 地球」在調查程序上的應用效果，有可能形成史上最有效益的虛擬清冊，再搭配調查員改為使用各種行動通訊工具，從而實現「行動調查法」。

### (九)數位電視機上盒法

在方法論的基礎上革新，根本取消「訪問程序」，而發展「數位行為調查法」。

數位電視機上盒法不僅會是一種調查工具，更有潛力成為人類行為科學實現控制、或潛移默化的工具，詳見吳統雄：『[老大哥悄悄的走近了：「機上盒」躡入的衝擊與瞻望](#)』專文。