抽樣方法

開課班級：應統二 1　　授課教師:郭清章

參考書籍：抽樣方法(2版) 黃文隆,黃龍(101) 滄海書局

抽樣調查的特點

抽樣調查(sampling survey)是市場調查中使用頻率最高的一種調查方式，它是按照一定程式，從所研究物件的全體（母體）中抽取一部分（樣本）進行調查或觀察，並在一定的條件下，運用數理統計的原理和方法，對母體的數量特徵進行估計和推斷。

與普查相比，抽樣調查具下三個顯著特點：

- 經濟 – 人力、物力、財力上
- 高效率 – 易滿足時效性要求
- 準確 – 參加人員較單純、品質好控制

抽樣調查的作用

對一些不可能或不必要進行普查的社會經濟現象，可用抽樣調查方式解決。例如，對有破壞性質或損耗性質的商品質量檢驗；對一些無限母體的調查(如對森林木材積蓄量調查)等。

在經費、人力、物力和時間有限的情況下，採用抽樣調查方式，可節省開支，爭取時效，用比較少的人力、物力和時間，達到滿意的調查效果。

抽樣方案設計

如何兼顧抽樣效果和所付出的代價，使之能有機會結合，一直是擺在抽樣方案設計者面前的一個重要問題，也是抽樣方案設計的難點所在。

抽樣方案設計的基本內容有：

- 第一、確定抽樣調查的目的、任務和要求
- 第二、確定調查物件(母體)的範圍和抽樣單位
- 第三、確定抽取樣本方法
- 第四、確定必要的樣本數
- 第五、對主要抽樣指標的精確度提出要求
- 第六、確定母體目標量的估算方法
- 第七、制訂實施母體方案的辦法和步驟

抽樣方案設計的主要步驟如圖所示：

定義母體及抽樣單位

↓

選擇抽樣架構(名冊)

↓

選擇抽樣技術

↓

確定必要的樣本

↓

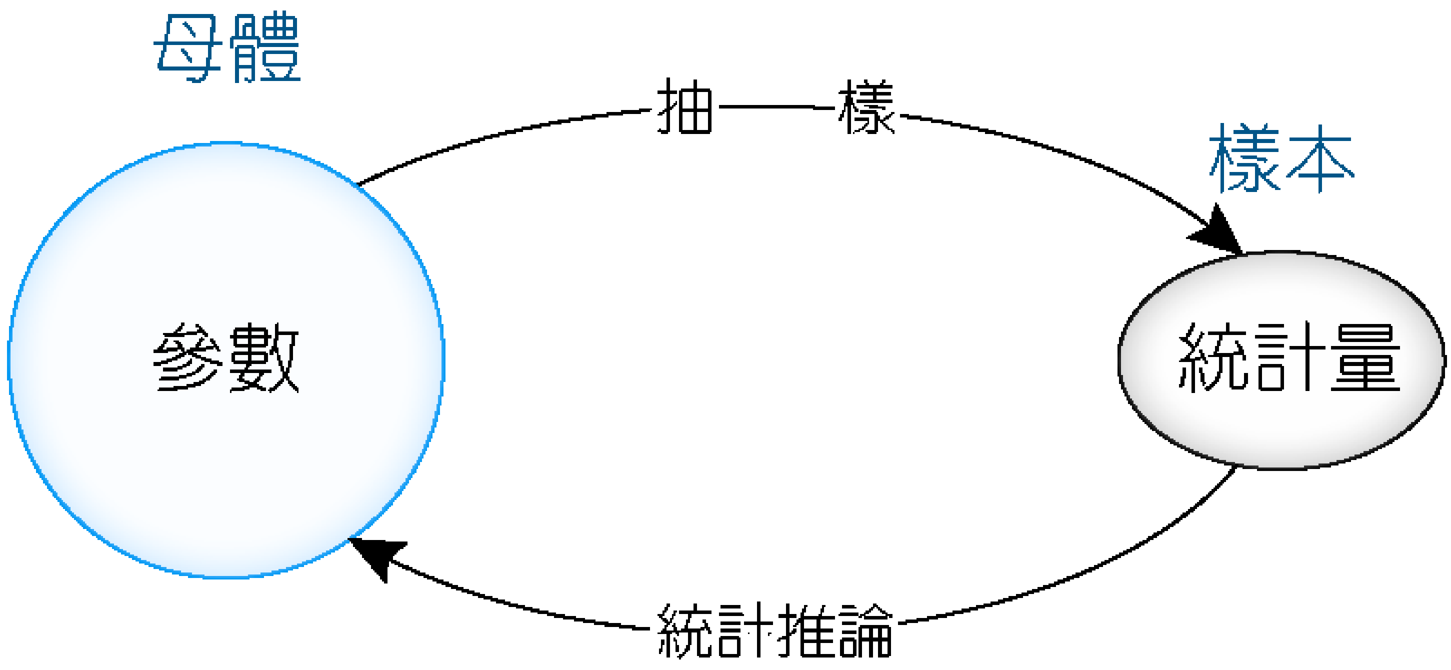
確定母體目標量的估量

↓

實施抽樣

抽樣的基本概念

- 「統計學」的核心課題是「統計推論」。而統計推論主要是藉由抽樣所得到的一組樣本的特性來推測或瞭解未知母體的主要特性。
- 所謂**抽樣(sampling)**，是運用適當的抽取方式，自母體中抽出部份樣本點的方法或程序。
- 抽樣的主要目的，是希望能藉由適當的抽樣方式，自母體中抽取一組具有代表性的樣本，並計算此樣本的特徵量數值（稱統計量statistic），再有效的、精確的推論母體的主要特徵量數（稱**參數**parameter）。
- 為了能有效的、精確的推論母體的主要特徵量數（**參數**），在抽樣時必須考慮採用正確而且合適的抽樣方法，才能獲得一組具有代表性的樣本。


母體
抽——樣
樣本
參數
統計量
統計推論

抽樣方法

- 為了取得具有代表性的樣本，在決定採用哪一種抽樣方法前，必須慎重考量統計資料的性質、分佈範圍，以及統計推論的目的。一般常用的抽樣方法可分為「非隨機抽樣」與「隨機抽樣」兩種。
  - 一、隨機抽樣法(random sampling)
    - 若抽樣之過程符合下列三個條件，則稱此種抽樣方法為隨機抽樣法：

1. 母體內任何一個元素均有被抽出的可能性。
2. 任一組樣本被抽取的機率均為已知或可計算得知。
3. 任一組樣本被抽取的過程均為獨立。即母體內某一個體被抽到與否不會影響其他個體被抽到之機會。

抽樣方法

- 二、非隨機抽樣法(non-random sampling)
  - 統計調查時所抽取的樣本並不是依照機率模式設計去抽取，而是根據個人的主觀意志；例如自己的專長、知識、研究的目的或考慮資料取得的方便性，來選取樣本的方法，稱為「非隨機抽樣法」。以「非隨機抽樣」所得到的樣本來推論母體特性，比較不具有公正性，故本節不擬討論非隨機抽樣的方法。

- 符合上述三個條件的抽樣稱為「**隨機抽樣**」。經由隨機抽樣得到的一組樣本，稱為**「隨機樣本」**(random sample)。
- 隨機抽樣依隨機樣本被抽出的機率不同，可分為「簡單隨機抽樣法」、「分層隨機抽樣法」、「群集隨機抽樣法」 、「多階段抽樣」以及「系統隨機抽樣法」等。

簡單隨機抽樣(simple random sampling)

- 從一個含有N個元素的母體中，隨機抽取數個（n個）元素（樣本點）為一組樣本，每一個樣本點被抽到的機會均相同，此種抽樣的方法，稱之為「**簡單隨機抽樣法**」；而按此種方法所抽出的樣本，則稱之為「**簡單隨機樣本**」，簡稱為「隨機樣本」。

- 在採用簡單隨機抽樣時，依其所抽取樣本點是採放回或不放回方式，又可分為「抽樣放回」(sampling with replacement) 與「抽樣不放回」(sampling without replacement)兩種方法。此兩種方法在抽樣時，樣本出現的機率並不相同。採取「抽樣**放回**」的簡單隨機抽樣，每組樣本出現的機率為 $(1/N)^n$，而採用「抽樣**不放回**」的簡單隨機抽樣時，其每組樣本出現的機率則為 $1/C_n^N$ 。

- 然而當母體個數相當大(N→∞)時，$(1/N)^n$與 $1/C_n^N$兩者差異已不大，此時兩種抽樣方法都可視為是獨立的狀況，也就是說不管樣本放回或是不放回，對於下一次抽取並不會造成影響。
- 另外依抽樣工具的不同，常用的簡單隨機抽樣方法有抽籤法或亂數表法等。
- **抽籤法**：將母體內N個元素加以編號，並將號碼放入箱內，由箱內任意抽取出n個籤碼，再由母體中抽取出與籤碼相符的n個樣本點，組成一組簡單隨機樣本。

- **亂數表法**：將母體N個元素加以編號，再依據亂數表（**如附錄**）上任意一號碼點為起始號碼，然後依序或隨意選取n個**亂數碼**，（注意：只選取屬於母體編碼範圍內的亂數碼），再由母體中挑取出與n個亂數碼相符的n個樣本點，組成一組簡單隨機樣本。

亂數表

	1	2	3	4	5	6	7	8	9	10
1	6824	7709	3937	3289	9545	0620	3904	5203	6590	8769
2	0237	7574	8607	1502	4776	0944	4946	1519	4834	2810
3	1336	8960	2192	7132	9267	4262	6070	7664	7690	3873
4	6840	3016	3991	8582	1813	0012	3781	8635	0286	3932
5	5577	7452	9477	7942	7328	0822	7876	6379	9014	6845
6	3495	3500	9497	8688	7764	0017	1221	5816	8840	8573
7	5163	5127	5955	7826	0982	3563	7783	1575	7738	9146
8	3746	5767	5137	3846	9113	3394	5172	3745	2574	5275
9	0596	6736	4273	7665	8229	6933	6510	0093	4091	4567
10	6553	4267	4071	3532	0593	3874	5368	5295	6303	2629

例

- 某公司舉辦年終餐會，會中有抽獎活動，公司提供一部汽車及若干個家電作為員工抽獎用。參加餐會的員工每人發給一張兩頭印有相同號碼的摸彩券，一半為存根聯，一半為抽獎聯。撕下其中的抽獎聯投入摸彩箱中攪拌均勻，再依序請公司各級主管抽出摸彩券，以決定家電與汽車的中獎人。這就是「抽取不放回的簡單隨機抽樣」。

分層隨機抽樣(stratified random sampling)

- 「分層隨機抽樣」是指抽樣前先將整個母體分成若干個(k個)不重疊的部份母體，稱之為「**層**」(strata)。分層時要盡量使層內各個體具有高度的同質性(homogeneity)，而層與層之間，則盡量要有明顯的差異性。然後再在各層內，按母體各層比例，以簡單隨機抽樣法抽出各層之隨機樣本 $n_{i,i=1,2,\ldots k}$ 。最後合併所抽到的各層隨機樣本($n_1+n_2+\ldots+n_k=n$)，即構成一組總樣本數為n的分層隨機樣本(stratified random sample)。

例

- 體育組想自全校同學中抽取100人做身高調查，為增加正確度，體育組採用分層隨機抽樣法抽樣，首先將全校5000個同學依身高（160公分以下、160～165公分、165～170公分、170公分以上）分為A,B,C,D四種層次，據統計全校5000個同學分成四層後每層人數分別為1300人、2000人、1200人、500人。佔全校人數比例分別為26%、40%、24%、10%。因此為了抽樣100人，可採用「比例配置法」，自A,B,C,D各層中分別隨機抽取26人、40人、24人、10人，總計100人，組成一組分層隨機樣本。

層別	母體(共 5000 人)100%	分層抽樣（n=100 人）
160 公分以下	第一層（1300 人）26%	第一層（隨機抽 26 人）
160～165	第二層（2000 人）40%	第二層（隨機抽 40 人）
165～170	第三層（1200 人）24%	第三層（隨機抽 24 人）
170 公分以上	第四層（500 人）10%	第四層（隨機抽 10 人）

群集隨機抽樣(Cluster random sampling)

- **「群集隨機抽樣法」(部落隨機抽樣法)**是將整個母體按地域關係或按方便性，分成若干個性質相似的「群集」或「部落」(cluster)。使得部落與部落之間同質性較高，而各個部落內的元素則彼此間的差異性較大。因而每一個部落均可視為整個母體的縮小。
- 群集隨機抽樣法：首先從事前分好的所有部落中，隨機抽取數個（個）部落為隨機部落樣本，第二步驟再對這些被抽到的隨機部落，作全面性的普查。

例

- 例如人口抽樣調查，可以每個家庭為一個（部落）單位，由所有家庭（部落）中隨機抽取數個（k個）家庭（部落），再從被選中的家庭作全部成員的調查。雖然採用部落抽樣法取得的樣本資料，有時會產生較大的誤差，然而因為可以就近集中調查，反而省下不少的調查時間與調查費用。故此種抽樣法還是有其可用之處。

例

- 原住民抽樣調查是典型的部落抽樣的例子，遍佈全台的原住民各部落之間，基本上都具有相似的原住民部落特性，故可以全台灣的原住民部落為抽樣單位，抽取數個（k個）部落（村），調查員再到被抽中的部落（村），做全面性的人口普查。所得的樣本資料，稱為「部落隨機樣本」。

系統抽樣(systematic sampling)

第一步驟：將母體內N個元素依序由1至 N 加以編碼。

第二步驟：再將母體內N個元素的編碼，平分為n個段落。

第三步驟：先由第一段落(1~ k)的編號中，隨機抽取一個號碼，（假設恰巧抽到2號），則在母體的第一段落(1~ k)的編號中，選取2號元素為第一個樣本點。

第四步驟：依第一個被抽到的樣本編號，每次加k個單位再取下一個樣本，直到編號超過母體個數，或直到選滿n個樣本為止。

$$(2,\ 2+k,\ 2+2k,\ 2+3k,\ \ldots,\ 2+(n-1)k)$$

例

- 例如N=110人，欲抽取人，則先將母體分成11個間段，每10號為一個間隔(k=10)。假設從第一段落編號為(1~10)中，隨機剛好抽中7號，依序繼續抽取第7+k號，第7+2k號，第7+3k號，…，第7+(11-1)k號。即有系統的依序抽出{ 7, 17, 27, 37, 47, 57, 67, 77, 87, 97, 107 }，共計11個元素為調查對象。

多階段抽樣(Multi-stage sampling)

多段抽樣是將選擇樣本的過程分成兩個或兩個以上的階段來完成。每一階段的抽樣，採取隨機抽樣的方式抽出該階段之抽樣單位。因此每一元素，被抽選的機率可由其在各階段被抽樣的機率相乘而得，故多段抽樣亦為一種機率抽樣

在抽樣調查中，母體常散佈甚廣，如採一段抽樣則費錢費時。因此在實際的大規模抽樣實務中，常採用多段抽樣。

例

- 以調查台南市國中一年級男生近視狀況為例，來說明多段抽樣的步驟。
  - (1).決定抽樣要分幾段進行，及每一段的抽樣架構為何。設分兩段，第一段抽樣架構是台南市所有的國中為抽樣單位，第二段抽樣架構是各國中之所有一年級班。
  - (2).決定各段的抽樣方法及樣本大小。第一段採取簡單隨機抽樣，抽出幾個國中。第二段採取集群抽樣，每一抽中的國中抽出兩個一年級班級。

週次	主題內容/章節	備註
第一週	課程介紹	
第二週	抽樣方法導論,統計概論	
第三週	抽樣方法導論,統計概論	
第四週	簡單隨機抽樣法	
第五週	簡單隨機抽樣法	
第六週	分層隨機抽樣法	
第七週	分層隨機抽樣法	
第八週	分層隨機抽樣法	
第九週	期中考	期中考
第十週	群集隨機抽樣法	
第十一週	群集隨機抽樣法	
第十二週	系統隨機抽樣法	
第十三週	系統隨機抽樣法	
第十四週	系統隨機抽樣法	
第十五週	比例估計法	抽樣問卷測驗報告
第十六週	多階段抽樣法	抽樣問卷測驗報告
第十七週	迴歸估計法	抽樣問卷測驗報告
第十八週	期末考	期末考

評分方式及比重

1.出席率、學習態度與平常測驗：30%

2.期中考：30%

3.期末考：40%