



# 第六章

## 抽樣設計

### (Sampling Design)

# 6.1 抽樣的本質

- 抽樣(Sampling)是自母體(Population)內選擇部分元素(Element)為調查對象。此調查對象稱為樣本(Sample)，相對是普查(Census)。
- 抽樣的假設是可經由樣本得知母體特性。
- 抽樣比普查更具經濟優勢，抽樣也較快速得到研究結果，有些情形不適用普查。
- 但樣本可能無法有效代表母體，而導致對母體推估的錯誤。
- 一般而言，樣本夠大就不會推估錯誤。

# 抽樣的重要性

- 較低的成本
- 較高正確性的結果
  - 抽樣研究的品質可能會高於普查研究，因為抽樣可使訪談(或施測)的品質提高，質性研究的深度訪談必然不可能普查。
  - 90%以上的研究誤差來自非抽樣誤差，只有10%或更少的比例來自抽樣誤差。
- 更快的得到研究成果
  - 產品品質抽測
- 母體成分之易於取得
  - 抽樣必須避免高估或低估母體特徵
- 抽樣相對於普查
  - 母體小、元素間差異大→普查

# 優良樣本的特性

- 正確性
  - 沒有偏誤，指樣本與母體特性相似的程度。
  - 正確樣本裡高估與低估的組合達成均衡，沒有系統變異。
  - 系統變異(**Systematic variance**)指測量方法的不當，導致測量結果產生偏離。
- 精準性
  - 估計的精準性(**precision of estimate**)
  - 抽樣誤差(**Sampling error**)
    - 抽樣過程會產生隨機變異，樣本特徵與母體特徵很難完全一致。
  - 估計標準誤(**Standard error of estimate**)
    - 衡量估計值與母體參數離散的程度
    - 估計標準誤愈小，代表樣本的精確程度愈高，愈能反應母體參數。
    - 抽樣首先必須考慮如何克服系統變異，並期望產生最小估計標準誤

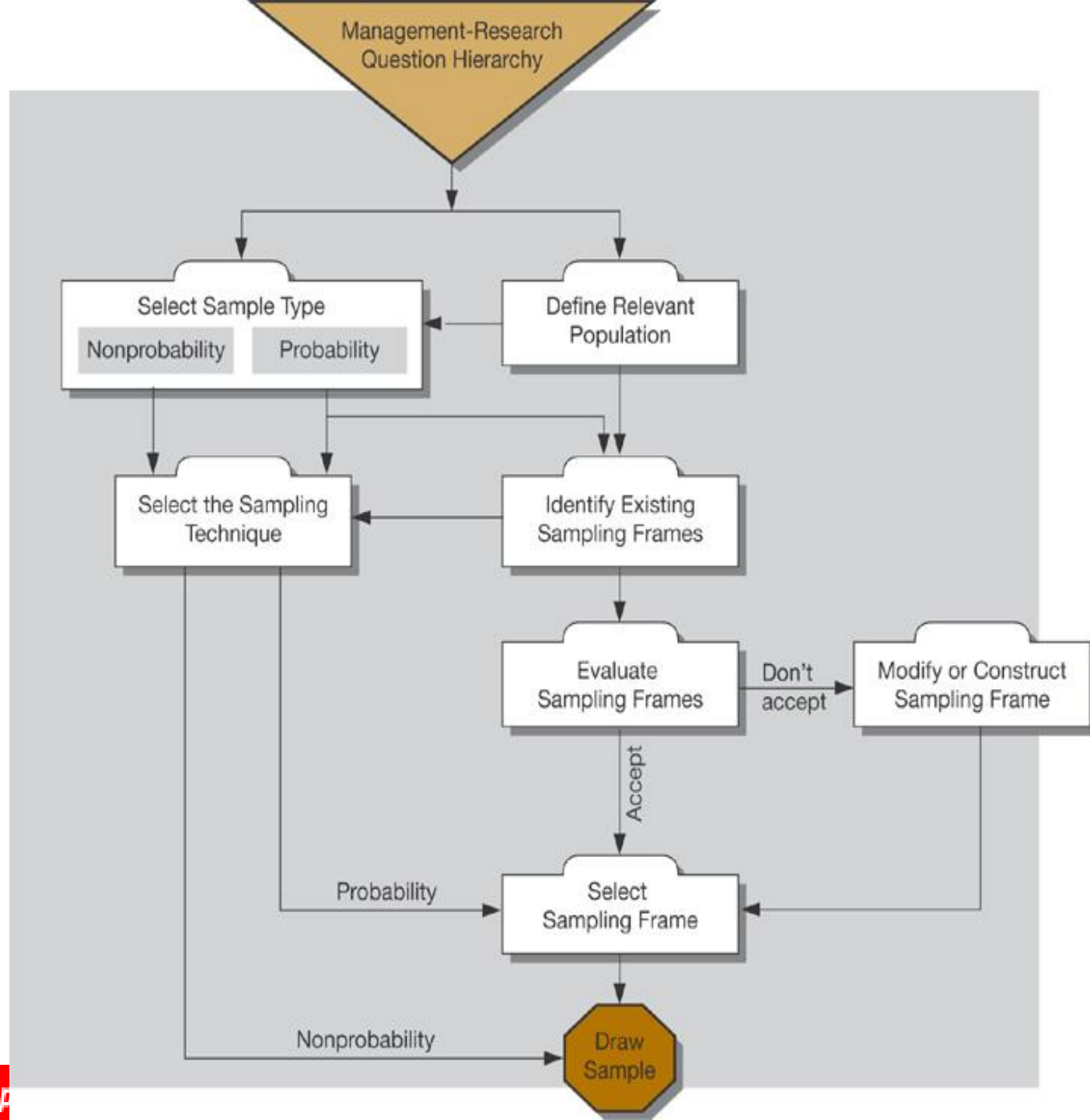
# 抽樣設計的類型(1/2)

- 抽樣的基本型態
  - 機率抽樣
    - 利用隨機選擇的概念，確保母體每一元素皆有被選中的機會。
  - 非機率抽樣
    - 主觀抽樣，無法確定每一元素都有被選中的機會。
- 樣本元素的選擇
  - 未限制抽樣
    - 每一個樣本元素皆以個體為單位從母體抽取。
  - 限制抽樣
    - 加上複雜法則先將母體分類之後，再進行抽樣。

# 抽樣設計的類型(2/2)

元素選擇 (Element Selection)	抽樣基本型態	
	機率型 (Probability)	非機率型 (Non-probability)
未限制抽樣 (Unrestricted)	簡單隨機抽樣 (Simple random)	簡便抽樣 (Convenience)
限制抽樣 (Restricted)	複雜隨機抽樣 (Complex random)	計畫抽樣 (Purposive)
	系統抽樣 (Systematic)	判斷抽樣 (Judgment)
	集群抽樣 (Cluster)	限額抽樣 (Quota)
	分層抽樣 (Stratified)	雪球抽樣 (Snowball)
	雙成抽樣 (Double)	

# 研究流程中之抽樣設計



## 6.2 機率抽樣

- 簡單隨機抽樣
  - 假設母體內每一元素被選擇機會皆已知且相等，每一元素都有已知非零的選擇機率。
- 抽樣設計的步驟
  1. 攸關母體為何？
    - 研究的母體到底為何？
  2. 關注的母體參數為何？
    - 指母體特質
  3. 抽樣主體(Sampling Frame)為何？
    - 抽樣主體必須近似母體清單，它應是正確完整的母體元素清單。
  4. 樣本型態為何？
    - 樣本的型態是否具有足夠的母體代表性。
  5. 所需樣本大小為何？
    - 樣本大小足以用來進行統計推論。
  6. 抽樣的成本多少？
    - 考量成本限制，例如使用電話、郵寄或親自訪談。



# 攸關母體的參數

- 母體
  - 母體為研究所欲觀察的對象集合
  - 研究者必須進行適當的操作型定義，將母體的結構、範圍定義出來。
- 母體參數
  - 對母體特性的描述值(如平均數、標準差、變異數)
- 樣本統計量
  - 對樣本特性的估計值
  - 樣本統計量可用來推估母體參數
- 資料型態
  - 名目、順序、區間、比率

# 抽樣主體(Sampling Frame)

- 抽樣主體應是正確完整的母體元素清單，至少必須相當近似母體清單。
- 抽樣主體常與理論母體有些差距
  - E.g. 某大學研究必須進行教職員生的調查，然而教職員生名錄可能每年十月份印製，如果研究進行時間是在五月，則可能因為人員異動使名錄與實際成員未能完全相符。
  - 上例中，抽樣主體即為教職員生名錄，理論母體則是全校教職員生。

# 樣本大小之計算

- 統計估計的誤差為  $Z(\sqrt{\frac{pq}{n}})$ ，所容許的誤差若為 **e**，則樣本大小之計算為

$$Z\sqrt{\frac{pq}{n}} \leq e \quad \text{交叉移項之後可得}$$

$$\sqrt{n} \geq Z \frac{\sqrt{pq}}{e} \quad \text{兩邊平方後得}$$

$$n \geq \frac{Z^2 pq}{e^2} \quad \text{n與變異程度成正比}$$

# 樣本大小之計算(cont.)

- 假設進行產品不良率研究，初步測試樣本不良率( $p$ )為10%，良率則為90%，若可容忍的誤差比率為3%，期望信賴程度為95.45%(即兩倍標準差， $Z=2$ )，帶入公式， $n=400$ ，即需要400樣本。
- 若事先未為不良率進行初測，則假定 $p=50\%$ ， $pq$ 乘積值最大值為0.25，其他條件不變，帶入上式， $n=1112$ ，此樣本數亦即為多數民意調查必須採樣的有效樣本數目。

- 附註：  
標準差公式 
$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

## 6.3 機率抽樣

- 機率抽樣方法的類型
  - 簡單隨機抽樣(Simple Random Sampling)
  - 分層抽樣法(Stratified Sampling)
    - 比率分層抽樣法
    - 非比率分層抽樣法
  - 集群抽樣法(Cluster Sampling)
  - 系統抽樣法(Systematic Sampling)
  - 雙重抽樣法(Double Sampling)

# 簡單隨機抽樣

- 簡單隨機抽樣
  - 自含有 $N$ 個個體之母體中，以不歸還的方式抽取 $n$ 個為樣本，每一個樣本被取出之機會相同，此種抽樣方法稱之。
- 共有  $C_n^N$  種組合。
- 簡單隨機抽樣設計在時間、金錢上都較為昂貴。

# 分層抽樣

- 分層隨機抽樣
  - 將母體中之全部個體分成若干不相重疊的部分，此部分稱為層(Strata)，在各層中分別以簡單隨機抽樣方式抽取樣本，集合各層之樣本為總樣本，此種方式稱為分層隨機抽樣。
- 要注意層的大小、分配比重、各層差異性與成本等。
- 比率分層抽樣
  - 各層樣本數佔全部樣本數比率與母體相同。
- 非比率分層抽樣
  - 依據主觀審慎判斷，確保每一層樣本大小足以產生充分的信心水準與區間估計。
  - 通常依據各層內母體數大小、內部變異性與抽樣成本而定。

# 分層抽樣(cont.)

- 分層抽樣的優點
  - 可提高抽樣精確度
  - 執行與管理上較為方便
  - 可在各層中分別求得估計值，研究者可深入探討不同分層樣本之間的差異。
- 實例
  - 在全縣各鄉鎮中之各村里按人口比例抽出若干合格選民集成本縣市全體樣本，以瞭解本縣市中選民之投票意向。
  - 在工廠裡，可將員工各生產線、機台、班別、部門別進行分層抽樣。



# 集群抽樣

- 集群抽樣法
  - 將全母體分成若干集群，再自集群中隨機抽取部分集群為樣本，將這些樣本之集合體當作總樣本進行調查，稱之。
- 集群與分層抽樣之差別
  - 分層抽樣以層內性質差異小、層間性質差異大為佳。
  - 集群抽樣之集群內各個體差異大，而集群間差異小者為佳。
  - 集群抽樣在選取某一集群後，即就該集群中之個體一一調查，可節省行政費用或旅費。
- 實例
  - 農工商業調查
    - 先畫分幾個區域別，再從區域別中抽出幾區進行調查研究。

# 集群抽樣(cont.)

分層抽樣法	集群抽樣法
母體被分成少數次群體，每個次群體所含元素數目相當多，其分類係依據與研究變數有關的特定標準。	母體被分成相當多個次群體，每個次群體所含元素少，其分類係依據資料蒐集時簡單易得的標準。
企圖使次群體內同質，而次群體間異質。	企圖使次群體內異質，而次群體間同質。
抽樣時從次群體內隨機選取元素。	抽樣時隨機選取數個次群體做為研究對象。

# 系統抽樣

- 系統抽樣
  - 在抽樣底冊中，隨機取一個單位為開頭，隨而每隔 $k$ 個單位取一單位，即構成「 $k$ 中取一」之系統樣本。
- 其中 $k \leq N/n$ ，以獲取總數 $n$ 個樣本。
- 系統抽樣易於瞭解、執行。
- 系統抽樣的單位成本較簡單抽樣法為低。

# 雙重抽樣

- 雙重抽樣
  - 將母體依據特定特徵分成若干次群體，用隨機方式抽出幾個次群體，再依據次群體所含訊息選擇次樣本之基礎。
- 類似分層抽樣或集群抽樣再進行隨機化。
- 又稱為序列抽樣法(sequential sampling)或多段抽樣法(multiphase sampling)
- 範例
  - 以調查學生參加俱樂部意願為例，先以電話或其他較為低成本的調查方法，探索哪些人願意參加俱樂部及其意願強度，根據意願強度分成若干層，再從各層中隨機抽樣做為進一步研究對象。

## 6.4 非機率抽樣

- 使用的理由
  - 可得到研究的目標
  - 較低的成本
  - 時間的限制
  - 不像完全隨機抽樣有那麼多的人為誤差
  - 無法取得全部的母體清單時

# 非機率抽樣方法的類型

- 簡便抽樣法(Convenience Sampling)
- 計畫抽樣法(Purposive Sampling)
  - 判斷抽樣法(Judgment Sampling)
  - 配額抽樣法(Quota Sampling)
- 滾雪球抽樣法(Snowball Sampling)

# 簡便抽樣

- 簡便抽樣
  - 研究人員自由選擇他們遇見的任何人
- 最便宜也最容易執行，但是最不可靠的抽樣方法。
- 探索性研究初期可以使用本法獲取某些基本訊息，其結果甚至可能推翻使用複雜抽樣方法的必要性。
  - 例如某研究員想要了解學生對某一議題的看法，他循序選取**25**名學生進行訪談，結果發現學生意見一面倒，這位研究員就不需要進行複雜抽樣的研究程序。
- 範例
  - 街頭攔截的意見調查。

# 計畫抽樣

- 計畫抽樣
  - 非隨機抽樣若是依據特定標準來選取樣本。
- 判斷抽樣法
  - 根據主觀設定某些標準的抽樣法。
  - 適合用在探索性研究初期，當研究人員希望選擇一個特定樣本做為篩選依據時，判斷抽樣法的可行性。
  - 範例
    - 企業經常以員工為樣本來試驗新產品，理由是員工對產品的評鑑能力優於一般大眾。
- 配額抽樣法
  - 遵照某些既定標準來抽樣。
  - 以某些重要特徵來描述母體屬性，若樣本的重要特徵與母體有相同分配，則在其他變數未被控制的情況之下，樣本也足以代表母體。
  - 範例
    - 廣泛用於民意調查、行銷研究。



# 滾雪球抽樣

- 滾雪球抽樣
  - 研究人員先以機率或非機率方式選出較容易辨識的個體元素成一組群，這個組群被用來尋找具有類似特徵的其他個體元素，形成新的組群，然後新的組群再被利用來辨識其他個體元素。
- 範例
  - 研究青少年幫派活動，事先並不知道哪些人是幫派份子，研究人員可從犯罪檔案找出特定份子，再以此為線索尋找其他幫派份子，直到取得足夠樣本為止。
  - 廣泛用於研究吸毒文化、幫派活動、政治權力菁英、社會關係與內部交易等。