

Review

統計方法的順序

1. 確定目的 (變量, 對象)
2. 蒐集資料 (方法: 普查, 抽樣)
3. 整理資料
4. 分析資料
5. 推論資料

1

第2章 抽 樣

- ◆ 2-1 母體與樣本
- ◆ 2-2 抽樣誤差與樣本大小
- ◆ 2-3 抽樣方法
- ◆ 2-4 隨機抽樣
- ◆ 2-5 非隨機抽樣

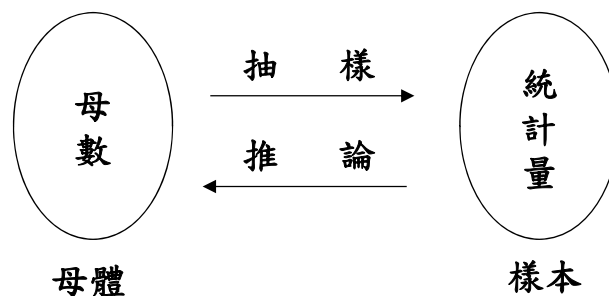
2

2-1 母體與樣本

蒐集統計資料的方法包括普查(Census)與抽樣(Sampling)兩種。而蒐集資料的全體對象稱為母體(Population)，對母體中每一個體作全面性的調查叫做普查，若是只抽出母體中的部分個體作調查就叫做抽樣，而被抽出的部分個體稱為樣本(Sample)，描述母體的特徵量數稱為母數或參數(Parameter)，描述樣本的特徵量數稱為統計量或表徵數(Statistic)。

3

樣本與母體的關係圖



4

■ 普查(Census)與抽樣(Sampling)

一般統計調查，大都使用抽樣，少用普查。原因是普查

- (1) 耗費過多人力、物力、時間、金錢等
- (2) 可能破壞母體

例如

- 了解湖泊中魚群的平均長度
- 燈泡耐用測試

5

■ 2-2 抽樣誤差與樣本大小

由抽樣產生的誤差稱為抽樣誤差，亦即是由樣本得到的統計量與母體參數值兩者間的差異。

凡是抽樣就不可避免會產生這種誤差，但是藉由增加樣本數，就會減少抽樣誤差。

但是基本上樣本數應以不少於30個較為理想。

6

■ 2-3 抽樣方法

抽樣的目的是在於獲得具代表性之樣本，以了解母體的情況，因此抽樣的方法非常重要。

抽樣的方法包括隨機抽樣 (Random Sampling) 與非隨機抽樣 (Nonrandom Sampling) 兩大類。

由隨機抽樣得來的樣本稱為隨機樣本 (Random Sample)，由非隨機抽樣得來的樣本稱為非隨機樣本 (Nonrandom Sample)

隨機樣本較客觀，適合用來做推論統計之用。非隨機樣本較主觀，不適合用來做推論統計之用。

7

■ 2-4 隨機抽樣

隨機抽樣又稱為機率抽樣 (Probability Sampling) 指的是母體中的每一個體都有可能被抽到，其方法是隨機的、非故意的

隨機抽樣一般可分為下列四種：

簡單隨機抽樣 (Simple Random Sampling)

系統抽樣 (Systematic Sampling)

分層抽樣 (Stratified Sampling)

集群抽樣 (Cluster Sampling)

8

■ 2-4-1 簡單隨機抽樣

將母體中的每一個體按照順序給予編號，再運用隨機性的方法，一次抽完所需要的樣本數，這樣的抽樣叫做簡單隨機抽樣。其可能利用的工具包括籤筒、彩券、號碼球與隨機號碼表(亂數表)。

如果母體內的個體數量不大，而且個體性質分布較均勻時，使用簡單隨機抽樣較理想。

9

■ 例題2.1

利用表2-1所示之隨機號碼表，以簡單隨機抽樣從班上40位同學中抽出6位當公差。

- 解**
- (1)編號：利用現有座號代替每一位同學（設座號為01~40號）。
 - (2)決定方向：先不看表，決定讀出編號的順序，例如先由左而右，次由上而下讀出編號。
 - (3)決定起點：先不看表，隨意決定以某行某列當起點，例如指定由第六行第七列當起點，則表2-1的起點為1。

10

■ 例題2.1



- (4)寫下編號：按照(2)之方向，取出與編號相同的位數之相關數字，以本題為例則需每兩位數一取，結果如下：13、54、66、14、14、98、46、25、19、34、55、05……。
- (5)決定樣本：捨棄超過40和重複的編號，則選出之六名公差為13、14、25、19、34與05號。

11

■ 隨機號碼表

91446 62290 40894 35102 84276 75127 45364 42346 60626 75492
 72865 13546 61414 98462 51934 55057 31621 81691 19236 52803
 24102 09223 64060 80718 63355 74414 21514 74940 85873 61390
 83563 47729 87309 37789 90455 44049 06678 30491 86017 02918
 28837 39763 67947 59462 58438 29629 03568 10493 27340 55201

12

■ 隨機號碼表

表 2-1 隨機號碼表的一部分

84652	94248	08376	01335	99248	52061	22135	55985	58709	13340
57881	90025	93613	48487	60454	36107	03492	90655	11782	72965
35946	14698	36089	12420	63673	53865	45293	00163	82618	23399
08686	16812	77721	19176	61667	09251	53843	94470	36355	36234
95480	45736	03002	24010	30380	51197	63459	69890	74588	46387
91446	62290	40894	35102	84276	75127	45364	42346	60626	75492
72865	13546	61414	98462	51934	55057	31621	81691	19236	52803
24102	09223	64060	80718	63355	74414	21514	74940	85873	61390
83563	47729	87309	37789	90455	44049	06678	30491	86017	02918
28837	39763	67947	59462	58438	29629	03568	10493	27340	55201

13

■ 2-4-2 系統抽樣

又稱等距抽樣，乃將母體中的每一個體按照順序給予編號，接下來再按照一定的距離，等間隔的抽完所需要的樣本數。如果母體內的個體數量較大，而且個體性質分佈較均勻時，使用系統抽樣較理想。

實施系統抽樣，首先要知道需間隔多遠抽一個樣本，這就是所謂的抽樣區間(Sampling Interval)，其求法如下：

$$\text{抽樣區間} = \frac{\text{母體個體總數}}{\text{樣本數}}$$

14

■ 例題2.2

以系統抽樣從班上40位同學中抽出6位當公差。

解

首先決定以四捨五入求出抽樣區間為 $40/6 \doteq 6.6 \doteq 7$ (人)，其次從01~07等座號區間以簡單隨機抽樣抽出第一人，假設抽出6號，接下來依序加上等間隔7，則抽出之公差為6、13、20、27、34、1等六人，請留意最後被抽出的人原為41號，因該班無此座號，在將座號首尾相接的情況下，因而選出1號。

15

■ 2-4-3 分層抽樣

將母體依照某種特性分成若干層(Strata)，再由各層隨機抽取所需樣本數的方法，稱為分層抽樣。

分層抽樣的使用時機是當母體中的個體分佈不均，且具有分層的傾向時用之，可避免造成抽出樣本的代表性不夠。

16

■ 例題2.3

銀行欲抽出500人了解該行信用卡持卡人的平均消費情況，已知持卡人計10,000人，可再分成高消費者2,000人，中消費者5,000人，低消費者3,000人，則該銀行應如何進行抽樣？

17

■ 例題2.3

解 因母體10,000人中，可分成三層性質差異大的消費族群，故採分層抽樣為宜。

此時高、中、低三層人數比例為

$$2000 : 5000 : 3000 = 2 : 5 : 3$$

故必須從高消費者以簡單隨機抽樣法抽去

$$500 \times 2 / (2+5+3) = 100 \text{ (人)}$$

從中消費者以簡單隨機抽樣法抽出

$$500 \times 5 / (2+5+3) = 250 \text{ (人)}$$

從低消費者以簡單隨機抽樣法抽出

$$500 \times 3 / (2+5+3) = 150 \text{ (人)}$$

合併三層所抽出的人數即是 $100+250+150=500$ (人)

18

■ 2-4-4 集群抽樣

又稱部落抽樣，就是將母體依某種標準分成若干差異甚小的群(Cluster)，而每一群就好像是母體的小縮影，接下來視樣本數的多寡，再隨機抽取一些群體，再對這些群體作全面的調查，稱為集群抽樣。

集群抽樣中群與群的性質差異小，但群內個體性質差異大，所以其分類標準恰與分層抽樣相反。如果母體內的個體性質分佈不均勻時，且具有組織化、集群化的現象時，使用集群抽樣較理想。

19

■ 例題 2.4

某高中一年級新生依常態編班成30班，每班50人，今欲從一年級新生抽出100人以了解身高情況，應如何進行抽樣？

解 在常態編班下此30班的同質性高，每個班如同所有新生的縮影，故可利用集群抽樣抽出其中兩個班，再對這兩個班的所有同學（合計100人）作全面的調查。

以上所述四種抽樣方法各有其適用特性，但有時母體十分龐大時為增加樣本的代表性，可以將上述四種方法混合使用。

20

■ 隨機抽樣方法之適用性

母體內的個體性質分佈不均勻，且具有組織化、集群化的現象

母體中的個體分佈不均，且具有分層的傾向

母體內的個體數量不大，而且個體性質分布較均勻

母體內的個體數量較大，而且個體性質分佈較均勻

21

■ 2-5 非隨機抽樣

非隨機抽樣又稱非機率抽樣

(Non-probability Sampling)

凡是抽樣時，母體中的每一個體被抽出的機率為不可知，則稱此種抽樣為非隨機抽樣。

這類抽樣往往靠著抽樣者個人的方便或判斷來抽取樣本，其方法頗為主觀，因此無法評估所選出的樣本是否具有代表性，也不適合推論統計之用。

22

■ 2-5-1 便利抽樣

又稱偶遇抽樣(Accidental Sampling)，樣本的選擇完全只考慮方便性，例如常見有人在車站作市場調查，受訪者必須是那時剛好經過該地點的人。

23

■ 2-5-2 判斷抽樣

又稱立意抽樣(Purposive Sampling)，樣本的選擇完全憑抽樣者的知識或經驗主觀地來取樣，其方法甚為簡單，但要小心人為偏見的發生。例如老師在課堂上想要抽出一位同學回答問題，故意叫某位成績很好的同學來回答，因為老師猜想該同學應回答得出來。

24

習題

1. 比較參數與統計量之間的差異。

答

2. 說明隨機抽樣較適合用來推估母體的理由。

答

25

習題

3. 指出下列各個例子分別屬於何種抽樣？

(1) 汽水工廠在生產線上每隔一段距離即抽出一瓶汽水，檢驗其品質。

答

(2) 統一發標利用搖獎機的號碼球開獎。

答

26

習題

(3) 學校指定各班班長代表該班參加師生座談會。

答

(4) 從某校所有男性學生與女性學生各依人數比例抽出一定人數，以了解全校學生身高情況。

答

27

習題

(5) 醫生利用自己的門診病人作樣本，觀察某種藥物的療效。

答

(6) 生物學家從某地區數十個蜜蜂窩中挑出五個，針對這五個蜜蜂窩的所有蜜蜂作調查，以了解該地區平均每個蜜蜂擁有多少隻蜜蜂。

答

28