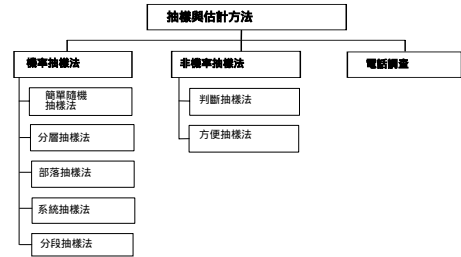


22 抽樣與估計方法

學習目的

1. 了解各種抽樣分法。
2. 了解不同抽樣方法的優缺點。
3. 了解各種抽樣方法之母體平均數、母體總和、母體比例與信賴區間之估計。
4. 利用 Excel 做統計估計。

本章結構



簡單隨機抽樣法

○ 簡單隨機抽樣法

簡單隨機抽樣是指母體內所有可能樣本其被抽出的機率均相等的抽樣方法。由簡單隨機抽樣法所得的樣本稱為簡單隨機樣本 (simple random sample)，簡稱為隨機樣本。

簡單隨機抽樣法

○ 母體平均數 $(1 - \alpha)$ 的信賴區間 - 無限母體

$$\bar{X} \pm Z_{\alpha/2} \sigma_x$$

其中 $\sigma_x = \frac{\sigma}{\sqrt{n}}$ (樣本平均數的標準差)。

若 σ 未知，以 $S_x = \frac{S}{\sqrt{n}}$ 估計 σ_x ，因此信賴區間為：

$$\bar{X} \pm Z_{\alpha/2} S_x$$

○ 母體總和的點估計式：簡單隨機抽樣

$$\hat{T} = N\bar{X}$$

簡單隨機抽樣法

○ σ_x 的估計式

$$S_x = \frac{NS_x}{\sqrt{n}}$$

式中：若母體無限 $S_x = \frac{S}{\sqrt{n}}$

若母體有限個數為 N ，則 $S_x = \frac{S}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ 。

○ 母體總和 $(1 - \alpha)$ 的信賴區間

① σ^2 已知 $\hat{T} \pm Z_{\alpha/2} \sigma_x = N\bar{X} \pm Z_{\alpha/2} N\sigma_x$

② σ^2 未知 $\hat{T} \pm Z_{\alpha/2} S_x = N\bar{X} \pm Z_{\alpha/2} NS_x$

簡單隨機抽樣法

○ 母體比例的估計

$1 - \alpha$ 信賴區間為：

$$\hat{p} \pm Z_{\alpha/2} S_p$$

式中 \hat{p} ：樣本比例， $S_p = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$

若為有限母體，而 $\frac{n}{N}$ 很小，則 $S_p = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ 。

簡單隨機抽樣法

○ 樣本數的決定

$$n = \frac{N \sigma^2}{(N-1)D + \sigma^2}$$

其中： N ：母體個數
 σ^2 ：母體變異數，假設已知，若未知則以 S^2 估計之。

$$D = \frac{d^2}{(Z_{\alpha/2})^2} \quad (\text{若估計母體均數})$$

$$= \frac{d^2}{(Z_{\alpha/2})^2 N^2} \quad (\text{若估計母體總和})$$

d ：估計誤差(或信賴區間長度的 1/2)

簡單隨機抽樣法

○ 母體比例樣本數的決定

$$n = \frac{Np(1-p)}{(N-1)D + p(1-p)}$$

其中： p ：母體比例，若母體比例未知，則以 \hat{p} 估計或以 $\frac{1}{2}$ 代之。

$$D = \frac{d^2}{(Z_{\alpha/2})^2}$$

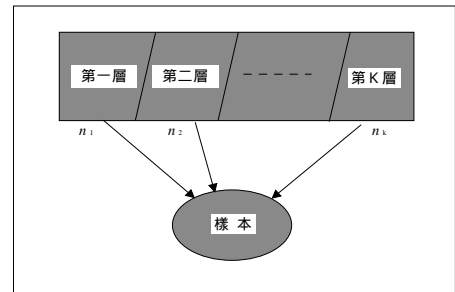
d ：估計誤差(或信賴區間長度的 1/2)

分層抽樣法

○ 分層隨機樣本

將母體依其特性分成不相重疊的組(稱為層)，自各層中簡單隨機抽取所得的樣本。

圖 22.1 分層抽樣方法



分層抽樣法

○ 母體平均數 μ 的估計式：分層抽樣

$$\bar{X}_c = \sum_{h=1}^H \left(\frac{N_h}{N} \right) \bar{X}_h$$

式中： H ：分層的組數

\bar{X}_h ：第 h 層的樣本平均數

N_h ：第 h 層的母體元素個數

N ：母體元素總個數， $N = N_1 + N_2 + \dots + N_H$ 。

分層抽樣法

○ 母體平均數 $(1-\alpha)$ 的信賴區間：分層抽樣

$$\bar{X}_c \pm Z_{\alpha/2} S_{\bar{X}_c}$$

分層抽樣法

- 母體總和的點估計式：分層抽樣

$$\hat{T} = N\bar{X}_s$$

- 母體總和估計式 \hat{T} 的估計標準差

$$S_{\hat{T}} = NS_s$$

- 母體總和 $(1-\alpha)$ 的信賴區間：分層抽樣

$$\hat{T} \pm Z_{\alpha/2} S_{\hat{T}} = N\bar{X}_s \pm NS_s$$

分層抽樣法

- 母體總和的點估計式：分層抽樣

$$\hat{T} = N\bar{X}_s$$

- 母體總和估計式 \hat{T} 的估計標準差

$$S_{\hat{T}} = NS_s$$

- 母體總和 $(1-\alpha)$ 的信賴區間：分層抽樣

$$\hat{T} \pm Z_{\alpha/2} S_{\hat{T}} = N\bar{X}_s \pm NS_s$$

分層抽樣法

- 分層樣本的母體比例的點估計式 (\hat{p}_s)

$$\hat{p}_s = \frac{\sum_{i=1}^h \left(\frac{N_i}{N}\right) \hat{p}_i$$

式中： h ：分層抽樣的層數， N_s ：第 h 層母體元素之個數， N ：母體元素的總個數， $N = N_1 + N_2 + \dots + N_h$ ， \hat{p}_s ：第 h 層樣本比例

- 母體比例 $(1-\alpha)$ 的信賴區間：分層抽樣

$$\hat{p}_s \pm Z_{\alpha/2} S_{\hat{p}_s}$$

分層抽樣法

- 比例配置法

$$n_s = n \left(\frac{N_s}{N} \right)$$

- 最優配置法

$$n_s = n \left(\frac{N_s \sigma_s}{\sum_{i=1}^h N_i \sigma_i} \right)$$

分層抽樣法

- 估計母體平均數時樣本數的決定：分層抽樣

$$n = \frac{\left(\sum_{i=1}^h N_i \sigma_i \right)^2}{N^2 \left(\frac{d^2}{Z_{\alpha/2}^2} \right) + \sum_{i=1}^h N_i \sigma_i^2}$$

式中：

N_s ： h 層的元素個數

N ：母體總個數

d ：估計誤差(或信賴區間長度的 1/2)

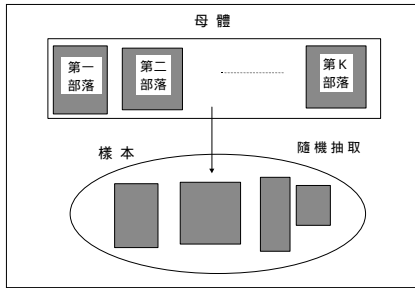
σ_s ： h 層的標準差，假設已知，若未知以 S_s 估計。

部落抽樣法

- 部落抽樣法

部落抽樣是先將母體中相鄰的某些群體劃分為 n 個不同的部落 (cluster)，母體中的每一個元素均屬於其中的一個部落，且是唯一的一個部落，然後再從這些部落中隨機抽取部落，並對抽出的部落進行普查的抽樣方法，又稱為集團抽樣。

圖 22.2 部落抽樣法



部落抽樣法

- 母體平均數的點估計式：部落抽樣

$$\bar{X}_c = \frac{\sum_{i=1}^K X_i}{\sum_{i=1}^K M_i}$$

- 母體平均數 $(1-\alpha)$ 的信賴區間：部落抽樣

$$\bar{X}_c \pm Z_{\alpha/2} S_{\bar{x}_c}$$

部落抽樣法

- 母體總和的點估計式：部落抽樣

$$\hat{T} = M\bar{X}_c$$

\hat{T} 的估計標準差為：

$$S_{\hat{T}} = MS_{\bar{x}_c}$$

- 母體總和 $(1-\alpha)$ 的信賴區間：部落抽樣

$$M\bar{X}_c \pm Z_{\alpha/2} MS_{\bar{x}_c}$$

部落抽樣法

- 母體比例的點估計式：部落抽樣

$$\hat{p}_c = \frac{\sum_{i=1}^K K_i}{\sum_{i=1}^K M_i}$$

K_i ：第 i 個部落，具某特性元素個數(或成功的次數) M_i ：第 i 個部落的元素個數

該點估計式估計標準差為：

$$S_{\hat{p}_c} = \sqrt{\frac{N-n}{NnM_c^2} \left(\frac{\sum_{i=1}^K (K_i - \hat{p}_c M_i)^2}{n-1} \right)}$$

- 母體比例 $(1-\alpha)$ 的信賴區間為：部落抽樣

$$\hat{p}_c \pm Z_{\alpha/2} S_{\hat{p}_c}$$

部落抽樣法

- 估計母體平均數時樣本數(部落個數)的決定：部落抽樣

$$n = \frac{N\sigma_c^2}{N \frac{d^2 M_c^2}{(Z_{\alpha/2})^2} + \sigma_c^2}$$

式中： σ_c^2 ：部落中元素的變異數 (variance of cluster of total X)

d ：估計誤差(或信賴區間長度的 1/2)

N ：母體中部落個數

M_c ：部落平均的元素個數 $\frac{M}{N}$

通常 σ_c^2 未知，因此以過去的資料來求其估計式：

$$S_c^2 = \frac{\sum (X_i - \bar{X}_c M_i)^2}{n-1}$$

部落抽樣法

- 估計母體比例時樣本數的決定：部落抽樣

$$n = \frac{N\sigma_c^2}{N \frac{d^2 M_c^2}{(Z_{\alpha/2})^2} + \sigma_c^2}$$

其中： σ_c^2 ：為部落中成功次數的變異數

d ：為估計誤差

通常 σ_c^2 未知，故求其估計式：

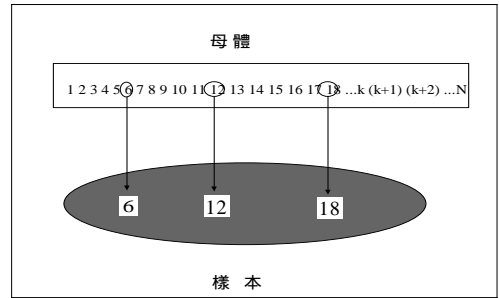
$$S_c^2 = \frac{\sum (K_i - \hat{p}_c M_i)^2}{n-1}$$

系統抽樣法

○ 系統抽樣法

系統抽樣法是自母體自然隨機排列的資料中，每隔一定間隔選取一個樣本，直至抽滿 n 個樣本為止。

圖 22.3 系統抽樣法

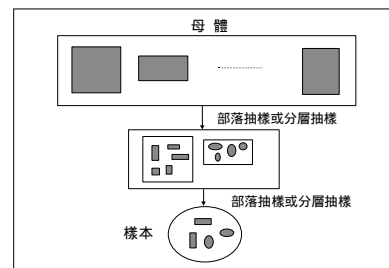


分段抽樣法

○ 分段抽樣法

分段抽樣是將母體分為數個階段，再以隨機或系統方法抽出樣本。

圖 22.5 分段抽樣法



判斷抽樣法

○ 判斷抽樣法

依自己的判斷來抽取樣本的方法稱為判斷抽樣法。

方便抽樣法

○ 方便抽樣法

以方便取得的方法來抽樣稱為方便抽樣法。