第七章　抽樣與抽樣分配

7.1 常見的抽樣方法

●隨機抽樣

1.簡單隨機抽樣法(Simple Random Sampling)

2.分層隨機抽樣法(Stratified Random Sampling)

3.部落抽樣法(Cluster Sampling)

4.系統抽樣法(Systematic Sampling)

●非隨機抽樣

1.便利抽樣

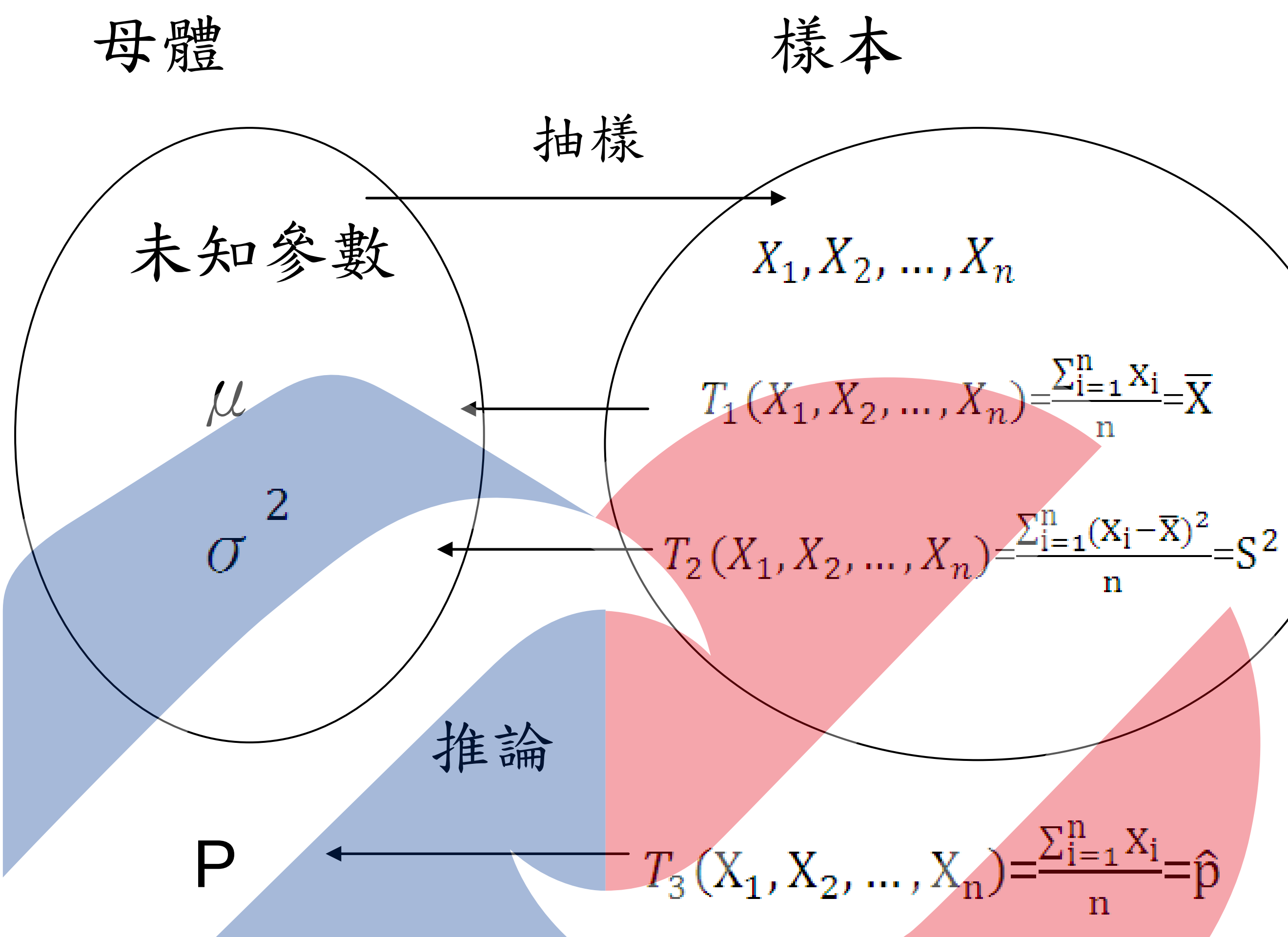
2.判斷抽樣

3.雪球抽樣

7.2 抽樣分配

抽樣分配：自一母體隨機抽取一樣本，組成實數值函數，該函數即為統計量(不含未知參數)，亦為隨機變數，其所形成的分配，謂之抽樣分配。

藉由抽樣分配連結統計量與未知參數，即探討抽樣分配的目的，是為了進行統計量推論(估計及假設檢定)未知參數作準備。



上述 μ, σ^2 及 p 稱為未知參數，$\bar{X}$, S^2 及 $\hat{p}$ 稱為統計量，接下來我們要探討 $\bar{X}$, S^2 及 $\hat{p}$ 的抽樣分配為何?

● $\bar{X}$的抽樣分配

1.放回抽樣

已知母體的平均數及變異數分別為μ及σ^2，自該母體以放回方式抽取 n 筆資料$X_1, X_2, \ldots, X_n$，則平均數$\bar{X}$之抽樣分配的期望值及變異數分別為 $E(\bar{X})=\mu$，$V(\bar{X})=\frac{\sigma^2}{n}$

證明：$E(\bar{X})= E(\frac{X_1+X_2+\cdots+X_n}{n})$

$=\frac{1}{n}[E(X_1)+E(X_2)+\ldots+E(X_n)]$

$=\frac{1}{n}n\mu$

$=\mu$

$V(\bar{X})= V(\frac{X_1+X_2+\cdots+X_n}{n})$

$=\frac{1}{n^2}[V(X_1)+ V(X_2)+\ldots+ V(X_n)]$

$=\frac{1}{n^2}n\sigma^2$

$=\frac{\sigma^2}{n}$

2.不放回抽樣

已知母體的平均數及變異數分別為μ及σ^2，自該母體以不放回方式抽取 n 筆資料$X_1, X_2, \ldots, X_n$，則平均數$\bar{X}$之抽樣分配的期望值及變異數分別為 E($\bar{X}$)=μ，V($\bar{X}$)=$\frac{\sigma^2}{n}\frac{N-n}{N-1}$

其中 N 表示母體個數，$\frac{N-n}{N-1}$稱為有限母體校正因子。

抽樣分配特性

1.當樣本數不大時，可以列出所有$\bar{X}$的可能值，進而求出$\bar{X}$的抽樣分配。

2.當樣本數很大時，不易求出$\bar{X}$的所有可能值，可藉由中央極限定理，獲得$\bar{X}$的近似抽樣分配。

3.實際抽樣皆以不放回方式抽樣，當$\frac{n}{N} \leq 0.05$時，$\frac{N-n}{N-1}$趨近1，此時已放回抽樣處理。

4.隨機樣本表示以樣本的隨機變數互為獨立且具有相同分配。

● 中央極限定理(central limit theorem)

若$X_1, X_2, \ldots, X_n$是從具有母體平均數μ與母體變異數σ^2的任何分配取出的隨機變數樣本，當 n 趨近於無限大時，

$\frac{\bar{X}-E(\bar{X})}{\sqrt{V(\bar{X})}} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$的累積分配函數會趨近於標準常態分配，即當 n 趨近無限大時，$\bar{X}$的分配趨近於常態分配 $N(\mu, \frac{\sigma^2}{n})$，應用上，當 n≥30 時，$\bar{X}$的抽樣分配就會近似常態分配 $N(\mu, \frac{\sigma^2}{n})$。