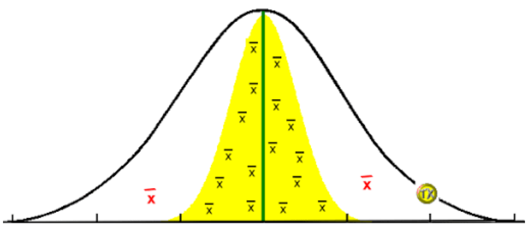
樣本平均數的抽樣分配

抽樣分配(sampling distribution)是統計推論的基礎，是很重要的觀念。在實務問題中，我們常常對母群體一無所知，就以隨機抽樣的方法從該母群體中抽出一組 n 個樣本，並以樣本的統計量來推論母群體的參數。由於我們只是從母群體中抽出了一組樣本，如果以同樣的方法再抽一組，甚至可以再抽好多好多組，結果雖然都不一樣，但是會有某種可預測的變異型態，抽樣分配就是描述這種型態。

以平均數為例，母群體的平均數是未知的，於是我們就從母群體(平均數為 μ ，標準差為 σ)中隨機抽出一組 n 個樣本，這組樣本平均數 $\bar{x}$ 是母群體平均數的一個估計值，如果我們以相同的方式，反覆的抽取 k 組樣本，每個樣本組都可以算出一個 $\bar{x}$，那麼 k 個 $\bar{x}$ 所形成的分配，就稱為平均數的抽樣分配。

理論上 $\bar{x}$ 抽樣分配的平均數會等於母群體的平均數 μ ，標準差則小於母群體的標準差 σ ，其關係如下圖所示：



有限母群體(不放回)：

有限母體抽出不放回樣本平均數的變異數 $(\sigma_{\bar{X}}^2)$ 與標準差 $(\sigma_{\bar{X}})$

$$\sigma_{\bar{X}}^2 = V(\bar{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$$

$$\sigma_{\bar{X}} = \sqrt{V(\bar{X})} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

無限母群體(放回)：

即 N 趨近於無限大，此時 (N-n)/(N-1) 就會趨近於 1，$\sqrt{(\mathrm{N}-\mathrm{n})/(\mathrm{N}-1)}$ 當然也趨近於 1，就可忽略了。$\sqrt{(\mathrm{N}-\mathrm{n})/(\mathrm{N}-1)}$ 一般稱為有限母群體修正項(finite population correction, fpc)。在實務上，母群體雖為有限，但如與樣本數比較起來相對很大時，此有限母群體修正項也會很小而可忽略不計。一般以樣本數等於或小於母群體的 5 %，即 n/N $\leq$ 0.05 時，雖來自有限母群體的樣本即可視為來自無限母群體。

樣本平均數的抽樣分配有以下的特性：

1. 每一組的樣本數 n 愈大時，其 $\bar{x}$ 會愈接近 μ 。

2. 樣本平均數 $\bar{x}$ 的平均數會等於母群體的平均數 μ 。

3. 母群體變異愈大，$\bar{x}$ 抽樣分配的彎異數也愈大。

4. 樣本平均數 $\bar{x}$ 的抽樣分配其變，異程度較母群體為小，且與樣本個數 n 成反比。當 n = 1 時，兩者之變異數完全相同，若 n 趨近於無限大時，$\sigma\ /\ \sqrt{n}$ 會趨近於 0 而不再有變異，亦即樣本平均數 $\bar{x}$ 的抽樣分配會收斂到母群體平均數 μ ，此稱為「大數法則(Law of Large Numbers) 。

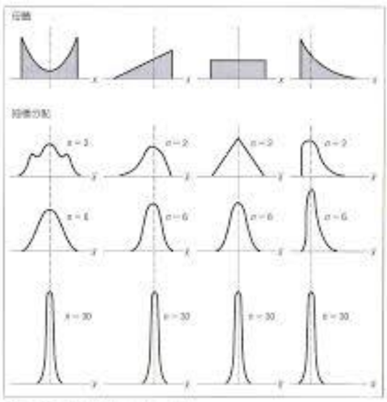


圖 8-2 各種母體應用中央極限定理的結果

5. 當 n 相當大時，儘管母群體之 x 不是常態分配，但其平均數的抽樣分配仍為常態分配，此即為中央極限定理。

6. 如樣本個數 n 占母群體之比例少於 5 % 或採收回抽樣時，雖為有限母群體，亦可使用無限母群體之公式。