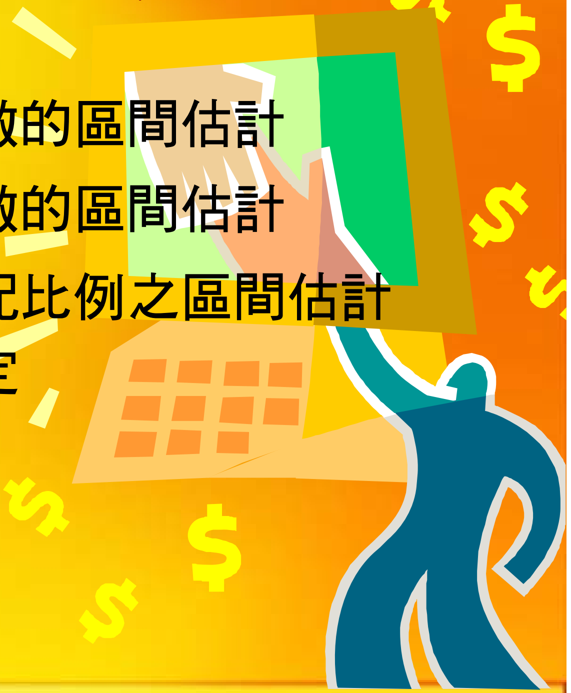
統計學

第八章 推論(一) － 估計



書號：512282

編著 江建良

普林斯頓國際有限公司

8-1 統計推論的概念

統計推論 (statistical inference)，即是由樣本推論母體的種種理論與方法。統計推論分為兩個步驟進行：

1. 先確立母體分配的型態
2. 對該母體分配進行未知母數的推論。

母體分配的型態：

連續性母體：

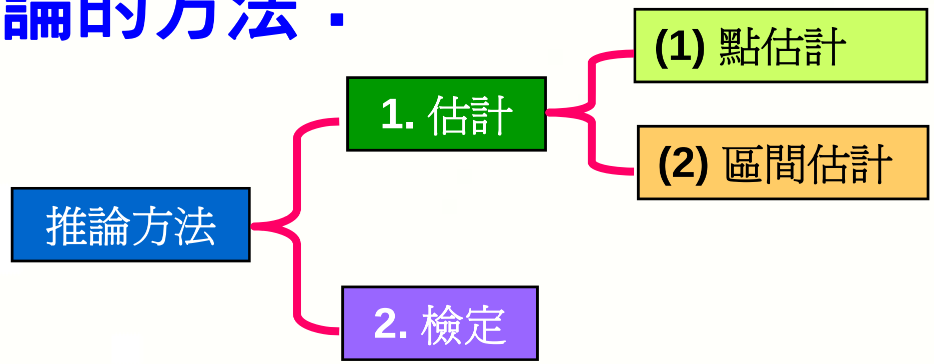
一般均假設為常態分配。

間斷性母體：

間斷性母體的分配型態，通常較易確定，
即所謂的點二項分配 。

8-1 統計推論的概念

統計推論的方法：



估計：

(1) 點估計：只根據樣本資料求算一個值來推估未知母數。

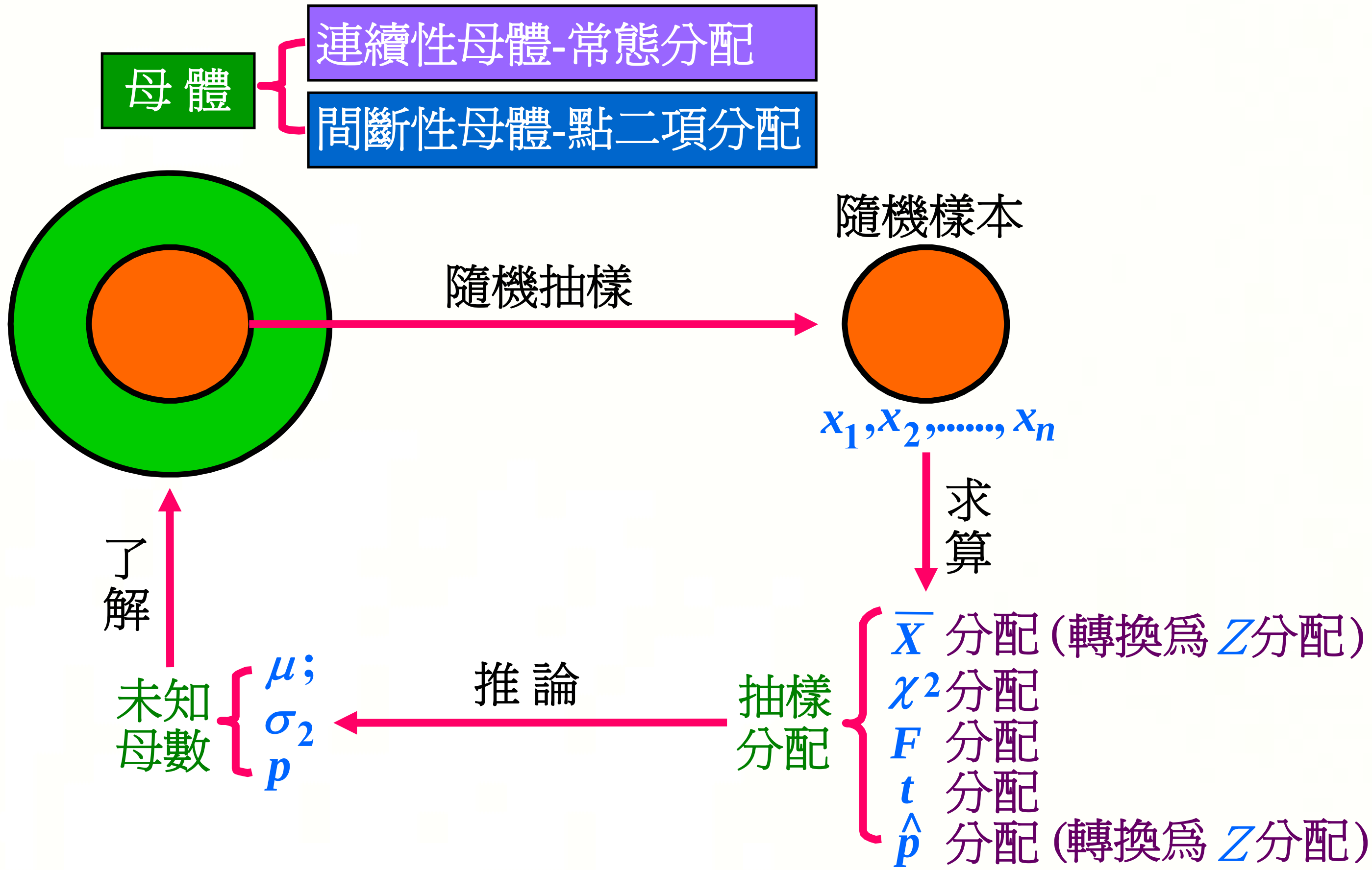
(2) 區間估計：不僅考慮點估計之值，更考慮推估可能產生的估計誤差。即對未知母數提供一個可能範圍的方法。

檢定：

對母體資料有適當的訊息，則可先對未知母數設一個假定值，再利用樣本資料的資訊，來測定假定值是否成立的方法。

8-1 統計推論的概念

- **統計推論的理論架構：**



8-2 點估計的意義與內容

點估計的意義：

即是利用樣本資料求算一個樣本統計量，用以推估母體中未知母數的方法。

點估計的內容：

母體未知母數 θ	點估計量 $\hat{\theta}$	與 $\hat{\theta}$ 有關之抽樣分配
母體平均數 μ	$\overline{X}$	Z，t
母體變異數 σ^2	$\hat{S}^2$	χ^2
兩母體平均數差 $\mu_1 - \mu_2$	$\overline{X}_1 - \overline{X}_2$	Z，t
兩母體變異數比 $\frac{\sigma_1^2}{\sigma_2^2}$	$\frac{\hat{S}_1^2}{\hat{S}_2^2}$	F
母體比例 P	$\hat{P}$	Z
兩母體比例差 $P_1 - P_2$	$\hat{P}_1 - \hat{P}_2$	Z

例題 8-1

今就常態母體中，抽取一組樣本，來估計該母體中未知的平均數，試以點估計的方式，說明之。

解：母體 $\xrightarrow{\text{抽樣}}$ 樣本 $\xrightarrow{\text{建立}}$ 統計量 $\xrightarrow{\text{即}}$ 估計量 $\xrightarrow{\text{求算}}$ 估計值 $\xrightarrow{\text{推估}}$ 未知母數

$$n(\mu, \sigma^2) \to x_1, x_2, \cdots, x_n \to \hat{\theta} = \overline{X} = \frac{1}{n}\Sigma X \to \qquad \to \hat{\theta} \to \mu$$

$$1,2,6,7 \longrightarrow \overline{X} = \frac{1+2+6+7}{4} = 4 \to \mu$$

（假設抽出 $n=4$ 的情況）

8-3 優良估計量的評判準則

不偏性：

係指點估計量 $\hat{\theta}$ 之期望值等於母體母數 θ，

即 $E(\hat{\theta})=\theta$，$\hat{\theta}$ 稱為 θ 的不偏估計量。

母體母數	不偏估計量
μ	$\overline{X}$
$\mu_1-\mu_2$	$\overline{X}_1-\overline{X}_2$
σ^2	$\hat{S}^2$
P	$\hat{P}$
P_1-P_2	$\hat{P}_1-\hat{P}_2$

例題 8-2

設 $X_i \sim n(\mu, \sigma^2)$，其中 $E(X_i)=\mu, V(X_i)=\sigma^2$，試判斷下列何者為 μ 的不偏估計量。

(a) $\hat{\theta}_1 = \dfrac{x_1+2x_2}{3}$ (b) $\hat{\theta}_2 = \dfrac{6x_1+4x_2}{10}$

(c) $\hat{\theta}_3 = \dfrac{x_1+4x_2}{4}$ (d) $\hat{\theta}_4 = \dfrac{x_1+x_2}{2}$

解：(a) $E(\hat{\theta}_1)=\dfrac{1}{3}[E(x_1)+2E(x_2)]=\dfrac{1}{3}\mu+\dfrac{2}{3}\mu=\mu$

(b) $E(\hat{\theta}_2)=\dfrac{1}{10}[6E(x_1)+4E(x_2)]=\dfrac{6}{10}\mu+\dfrac{4}{10}\mu=\mu$

(c) $E(\hat{\theta}_3)=\dfrac{1}{4}[E(x_1)+4E(x_2)]=\dfrac{1}{4}\mu+\mu=\dfrac{5}{4}\mu\neq\mu$

(d) $E(\hat{\theta}_4)=\dfrac{1}{2}[E(x_1)+E(x_2)]=\dfrac{1}{2}\mu+\dfrac{1}{2}\mu=\mu$

$\therefore$ (a), (b), (d) 是不偏估計量，(c) 不是。

由上例可知，只要 $\hat{\theta}=\alpha\hat{\theta}_1+(1-\alpha)\hat{\theta}_2$, $0<\alpha<1$, $E(\hat{\theta}_1)=E(\hat{\theta}_2)=\theta$，則 $\hat{\theta}$ 為 θ 的不偏估計量 [請讀者自行證明]。

8-3 優良估計量的評判準則

- **有效性：**

係指在樣本大小 n 固定之下，母數 θ 的所有不偏估計量 $\hat{\theta}$ 中具有最小變異數者，稱為有效不偏估計量。

- **相對有效性：**

$$\text{R.E.} = \frac{V(\hat{\theta}_1)}{V(\hat{\theta}_2)}$$

若 $\text{R.E.} < 1$，則 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效

$\text{R.E.} > 1$，則 $\hat{\theta}_2$ 比 $\hat{\theta}_1$ 有效

例題 8-3

試就上例說明何者為有效推定量？(設樣本大小 n 固定)

解：∵有效推定量的前提是，樣本大小 n 固定且 $\hat{\theta}$ 爲 θ 的不偏估計量，所以以 (a)(b)(d) 比較之。

$$V(\hat{\theta}_1)=V\left(\frac{x_1+2x_2}{3}\right)=\frac{5}{9}\sigma^2$$

$$V(\hat{\theta}_2)=V\left(\frac{6x_1+4x_2}{10}\right)=\frac{52}{100}\sigma^2$$

$$V(\hat{\theta}_4)=V\left(\frac{x_1+x_2}{2}\right)=\frac{2}{4}\sigma^2$$

其中 $\hat{\theta}_4$ 最小，所以 $\hat{\theta}_4$ 是三個估計量中最有效的。

例題 8-4

設 $X_i \sim n(\mu,\sigma^2)$，則樣本平均數 $\overline{X}$ 與樣本中位數 M_e，那一個是母體平均數 μ 的優良估計量？已知 $E(M_e)=\mu, V(M_e)=\frac{\pi}{2}\frac{\sigma^2}{2}=1.57\frac{\sigma^2}{n}$，設樣本大小 n 固定。

解： $\because X_i \sim n(\mu,\sigma^2) \quad \therefore \overline{X} \sim n\left(\mu,\frac{\sigma^2}{n}\right)$

已知 $M_e \sim n\left(\mu, 1.57\frac{\sigma^2}{n}\right) \quad \therefore E(\overline{X})=E(M_e)=\mu$

由 $\text{R.E.}=\dfrac{V(\overline{X})}{V(M_e)}=\dfrac{\frac{\sigma^2}{n}}{\frac{1.57\sigma^2}{n}}=\dfrac{1}{1.57}<1$

$\therefore \overline{X}$ 比 M_e 相對有效。

8-4 區間估計的概念

區間估計的意義：

係指根據抽出的樣本資料，先求出未知母數的點估計量(即樣本估計量)，再將點估計量作成抽樣分配，然後利用此抽樣分配建立一個未知母數的可能範圍之方法，稱為區間估計。

信賴區間的導出過程：

(1) 機率區間：

$$P(-Z_{(1-\frac{\alpha}{2})} \le Z \le Z_{(1-\frac{\alpha}{2})}) = P\left(-1.96 \le \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le 1.96\right) = 1-\alpha = 0.95$$

(2) 信賴區間：

$$\overline{X} - 1.96\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + 1.96\frac{\sigma}{\sqrt{n}}$$

例題 8-5

設 $X_i \sim n(\mu,\sigma^2)$，σ^2 為已知數，自其中隨機抽出 n 個樣本，令為 $x_1, x_2, \cdots, x_n$，設信賴水準 $1-\alpha=0.95$，試求母數平均數 μ 的信賴區間。

解：(1) 母體分配：常態分配 $f(x)=n(\mu,\sigma^2)$, σ^2 爲已知。

(2) 抽樣： $x_1, x_2, \cdots, x_n$ 之一組隨機樣本。

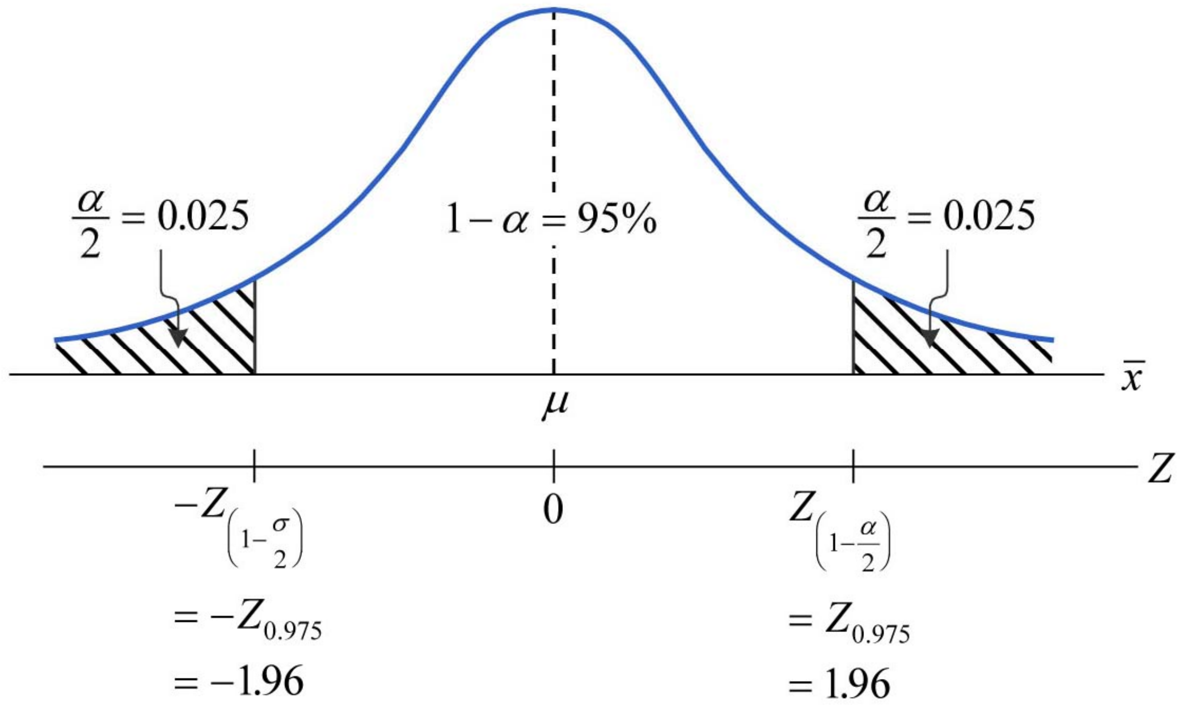
(3) 點估計量：以 $\hat{\mu}=\overline{X}$ 表示（設點估計量爲不偏估計量）。

(4) 樣本統計量：將點估計量 $\overline{X}$，轉換爲標準常態值

$$Z=\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

(5) 抽樣分配：Z 的抽樣分配，稱 Z 分配爲標準常態分配

$$\overline{X} \sim n\left(\mu , \frac{\sigma^2}{n}\right) \xrightarrow{\text{標準化}} z \sim n(0 , 1)$$



(6) 機率區間：Z 之機率區間為

$$P\left(-Z_{\left(1-\frac{\alpha}{2}\right)} \le Z \le Z_{\left(1-\frac{\alpha}{2}\right)}\right) = P\left(-1.96 \le \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le 1.96\right) = 1-\alpha = 0.95$$

(7) 信賴區間：因為我們的目的是求得 μ 之信賴區間，其型態為 $L \le \mu \le U$ ，因此必須對機率區間加以轉換，得

$$P\left(\overline{X}-1.96\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X}+1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

[將機率區間各項乘以 $\frac{\sigma}{\sqrt{n}}$ ，再減 $\overline{X}$ ，再乘以 (-1)]

(8) 結論：信賴區間 $L \le \mu \le U$

$$\Rightarrow \quad \bar{X} - 1.96\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow \begin{cases} \text{信賴區間的下限 } \bar{X} - 1.96\dfrac{\sigma}{\sqrt{n}} \\ \text{信賴區間的上限 } \bar{X} + 1.96\dfrac{\sigma}{\sqrt{n}} \end{cases}$$

可記為 μ 的信賴區間

$$\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right)$$

或 $\bar{X} \mp 1.96\frac{\sigma}{\sqrt{n}}$ （“－”表示減，“＋”表示加）

上式中，$\bar{X}$ 為點估計量，$1.96\frac{\sigma}{\sqrt{n}}$ 是為抽樣誤差(或稱誤差界限)，

因此我們亦可以下列通式來表示母體母數的信賴區間，即

母體母數之信賴區間＝點估計量 $\mp$ 抽樣誤差

8-4 區間估計的概念

- 機率區間與估賴區間的差異：

  - 機率區間是一個固定區間：

$$P\left(-1.96 \le \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le 1.96\right) = P\left(\underbrace{\mu - 1.96\frac{\sigma}{\sqrt{n}}}_{(\text{固定數})} \le \underset{(\text{隨機變數})}{\bar{X}} \le \underbrace{\mu + 1.96\frac{\sigma}{\sqrt{n}}}_{(\text{固定數})}\right) = 0.95$$

  - 機率區間是一個固定區間：

$$P\left(\underbrace{\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}}_{(\text{隨機變數})} \le \underset{(\text{固定數})}{\mu} \le \underbrace{\bar{X} + 1.96\frac{\sigma}{\sqrt{n}}}_{(\text{隨機變數})}\right) = 0.95$$

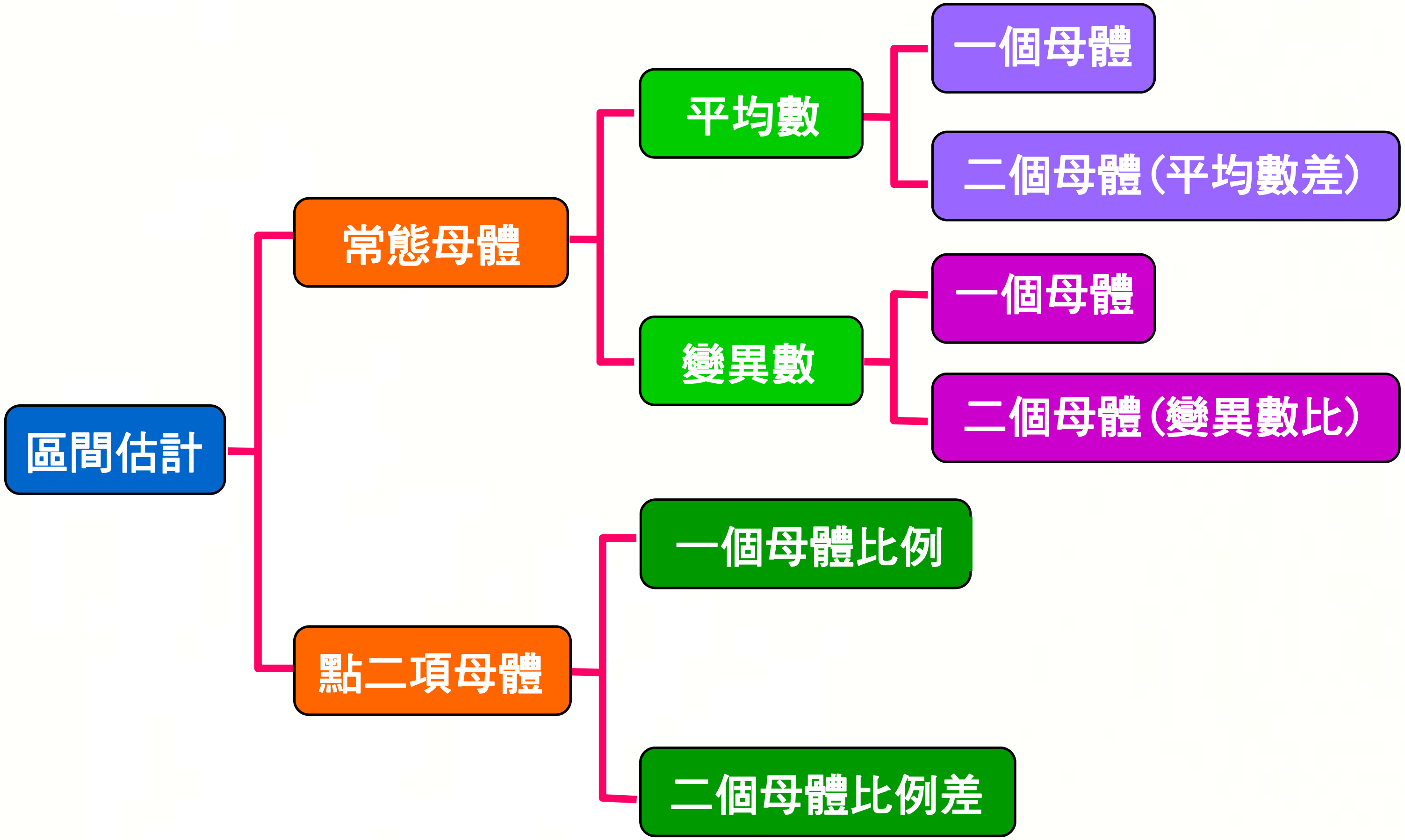
8-4 區間估計的概念

- **影響信賴區間大小的因素：**
  - 點估計量
  - 樣本大小 n
  - 信賴區間上下限取法
  - 信賴係數

- **優良信賴區間的衡量：**

1. 當信賴係數相同之下，以信賴區間愈短者愈精準，當兩尾各取 $\alpha/2$ 為最短。
2. 當信賴區間長度相同之下，以信賴係數愈大者愈精準。

8-5 常態母體平均數的區間估計



P264

8-5 常態母體平均數的區間估計

一個母體平均數的區間估計：

母體分配	母體標準差	樣本大小	抽樣分配
常態	已知	大	*Z*
常態	已知	小	*Z*
常態	未知	大	*Z*
常態	未知	小	*t*
未知	已知	大	*Z*
未知	已知	小	—
未知	未知	大	*Z*
未知	未知	小	—

註：大樣本係指 $n \geq 30$

1. 母體標準差已知，母體平均數 μ 的信賴區間：

$$\mu \sim \overline{X} \mp Z_{(1-\frac{\alpha}{2})} \cdot \frac{\sigma}{\sqrt{n}}$$

2. 母體標準差未知，小樣本，母體平均數的信賴區間：

$$\mu \sim \overline{X} \mp t_{(1-\frac{\alpha}{2},\, n-1)} \frac{\hat{S}}{\sqrt{n}}$$

例題 8-6

設自動販賣機銷售大杯可樂，其重量近於常態分配，其標準差為 2 公克，今抽取 100 杯為一樣本，其平均重量為 30 公克，試求：

(a) 該自動販賣機可樂平均重量 95% 之信賴區間。

(b) 該自動販賣機可樂平均重量 95% 之信賴區間長度。

(c) 該自動販賣機可樂平均重量 95% 之抽樣誤差（即誤差界限）。

解：(a) $n = 100$，$\sigma = 2$，$X_{(1-\frac{\alpha}{2})} = Z_{0.975} = 1.96$

$$\overline{X} - Z_{(1-\frac{\alpha}{2})}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + Z_{(1-\frac{\alpha}{2})}\frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow \quad 30 - 1.96 \cdot \frac{2}{\sqrt{100}} \le \mu \le 30 + 1.96 \cdot \frac{2}{\sqrt{100}}$$

$$\Rightarrow \quad 29.61 \le \mu \le 30.39 \text{（公克）}$$

(d) $30.39 - 29.61 = 0.78$（公克）

(c) 抽樣誤差 $= Z_{(1-\frac{\alpha}{2})} \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{2}{\sqrt{100}} = 0.39$（公克）

例題 8-7

東亞日光燈壽命呈常態分配，其變異數為 100 小時 ^{2}，今隨機抽取 $n=30$ 支日光燈，其平均壽命為 800 小時，試求東亞日光燈之平均壽命之 96% 的信賴區間。

解： $n=30$，$\sigma^2=100$，$\sigma=10$，$1-\alpha=0.96$，$Z_{(1-\frac{\alpha}{2})}=Z_{0.98}\doteq 2.05$

$$\mu \sim \bar{X} \mp Z_{(1-\frac{\alpha}{2})}\frac{\sigma}{\sqrt{n}} = 800 \mp 2.05\cdot\frac{10}{\sqrt{30}} = 800 \mp 3.74$$

$\therefore$ 求得 96% 之信賴區間爲 $796.26 \le \mu \le 803.74$（小時）

例題 8-8

自一箱蘋果中抽出 400 個，計算其平均重量 $=50$ 公克，標準差 8 公克，在 95% 之信賴係數下，試求：

(a) 整箱蘋果平均重量的估計值

(b) 整箱蘋果平均重量的信賴區間

解： 本題知母體分配未知，標準差未知，但 $n \geq 30$ ，利用中央極限定理，可使用 Z 分配求解。

(a) 點估計值 $= \overline{X} = 50$ 公克。

(b) $n = 400$ ， $\overline{X} = 50$ ， $\hat{S} = 8$ ， $1 - \alpha = 0.95$ ， $Z_{\left(1-\frac{\alpha}{2}\right)} = Z_{0.975} = 1.96$

$$\mu \sim \overline{X} \mp Z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}}$$

$$= \overline{X} \mp Z_{\left(1-\frac{\alpha}{2}\right)} \cdot \frac{\hat{S}}{\sqrt{n}}$$ (σ 未知，以 $\hat{S}$ 代之)

$$= 50 \mp 1.96 \cdot \frac{8}{\sqrt{400}} = 50 \mp 0.78$$

$\therefore\ 49.22 \leq \mu \leq 50.78$ (公克)

例題 8-9

今在甲學校隨機抽取 $n=16$ 名學生為樣本，調查其中餐費用平均 30 元，標準差 5 元，假設甲學校學生中餐費用為常態分配，試求該學校學生中餐平均費用之 90% 的信賴區間。

解： $\because n=16<30$ 爲小樣本，且母體標準差未知，而母體爲常態分配，故可使用 t 分配求解。

$n=16$ ，$\overline{X}=30$ ，$\hat{S}=5$ ，$f=n-1=15$ ，$t_{(1-\frac{\alpha}{2},n-1)}=t_{(0.95,15)}=1.753$

$$\mu \sim \overline{X} \mp t_{(0.95,15)} \cdot \frac{\hat{S}}{\sqrt{n}} = 30 \mp 1.753 \cdot \frac{5}{\sqrt{16}} = 30 \mp 2.19$$

$\therefore 27.81 \le \mu \le 32.19$ (元)

例題 8-10

設一箱雞蛋重量呈常態分配，自其中隨機抽取 10 個雞蛋，其重量分別為 10.2, 9.7, 10.3, 10.1, 10.1, 10.4, 10.3, 9.8, 9.9, 9.8 (公克)，試求該箱雞蛋平均重量之 99% 之信賴區間。

解：

$$\bar{X} = \frac{\Sigma X}{n} = \frac{10.2+9.7+10.3+10.1+10.1+10.4+10.3+9.8+9.9+9.8}{10}$$

$$= \frac{100.6}{10} = 10.06(\text{公克})$$

$$\hat{S} = \sqrt{\frac{\Sigma(X-\bar{X})^2}{n-1}} = \sqrt{\frac{(10.2-10.06)^2+(9.7-10.06)^2+\cdots+(9.8-10.06)^2}{10-1}}$$

$$= 0.246$$

$$t_{(1-\frac{\alpha}{2},n-1)} = t_{(0.995,9)} = 3.25$$

$$\mu \sim \bar{X} \mp t_{(0.995,9)} \cdot \frac{\hat{S}}{\sqrt{n}} = 10.06 \mp 3.25 \cdot \frac{0.246}{\sqrt{10}} = 10.06 \mp 0.25$$

$$\therefore 9.81 \le \mu \le 10.31(\text{公克})$$

8-5 常態母體平均數的區間估計

兩個母體平均數差的區間估計：

母體分配	母體標準差	樣本大小	抽樣分配
兩個常態母体	已知	大	Z
兩個常態母體	已知	小	Z
兩個常態母體	未知，但 1. $\sigma_1^2=\sigma_2^2=\sigma^2$	小	t
	2. $\sigma_1^2\neq\sigma_2^2$	小	t
未知	未知	大	Z

註：大樣本係指 $n_1 \geq 30$，$n_2 \geq 30$，母體標準差未知，以樣本標準差替代

8-5 常態母體平均數的區間估計

兩個母體平均數差的區間估計：

1. 兩個母體標準差已知，兩母體平均數 $\mu_1 - \mu_2$ 之信賴區間：

$$\mu_1 - \mu_2 \sim (\overline{X}_1 - \overline{X}_2) \mp Z_{(1-\frac{\alpha}{2})} \times \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

2. 兩母體標準差未知但 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 且小樣本，兩母體平均數 $\mu_1 - \mu_2$ 之信賴區間：

$$\mu_1 - \mu_2 \sim (\overline{X}_1 - \overline{X}_2) \mp t_{(1-\frac{\alpha}{2},\, f)} \times S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

3. 兩母體標準差未知，且不等 $\sigma_1^2 \neq \sigma_2^2$ 兩獨立樣本為小樣本，兩母體平均數 $\mu_1 - \mu_2$ 之信賴區間：

$$\mu_1 - \mu_2 \sim (\overline{X}_1 - \overline{X}_2) \mp t_{(1-\frac{\alpha}{2},\, f)} \times \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}}$$

4. 兩母體為常態分配，變異數未知，且兩小樣本不獨立(即成對抽取)，其母體平均數差 $\mu_1 - \mu_2$ 之信賴區間：

$$\mu_D \sim \overline{D} \mp t_{(1-\frac{\alpha}{2},\, f)} \frac{\hat{S}_D}{\sqrt{n}}$$

例題 8-11

甲班統計學成績標準差 = 5 分，乙班則標準差 = 3 分，今自甲乙兩班分別隨機抽出 $n_1=25$ 及 $n_2=36$ 名學生，計算其平均成績分別為 $\overline{X}_1=80, \overline{X}_2=75$ 分，假設兩班分數均成常態分配，試求甲乙班統計學成績平均分數差（甲班減乙班）之 94% 的信賴區間。

解： ∵為常態分配，標準差為已知，故可用 Z 分配求解。

甲班：$n_1=25, \overline{X}_1=80, \sigma_1=5$

乙班：$n_2=36, \overline{X}_2=75, \sigma_2=3$

$1-\alpha=0.94, Z_{(1-\frac{\alpha}{2})}=Z_{0.97}=1.88$

$$(\overline{X}_1-\overline{X}_2)-Z_{0.97}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \le \mu_1-\mu_2 \le (\overline{X}_1-\overline{X}_2)+Z_{0.97}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

$$(80-75)-1.88\sqrt{\frac{5^2}{25}+\frac{3^2}{36}} \le \mu_1-\mu_2 \le (80-75)+1.88\sqrt{\frac{5^2}{25}+\frac{3^2}{36}}$$

$2.898 \le \mu_1-\mu_2 \le 7.102$（分）

例題 8-12

自甲公司隨機抽取 $n_1 = 200$ 人，得知月薪平均 26,000 元，標準差 = 1,400 元，另自乙公司隨機抽取 $n_2 = 100$ 人，得知月薪平均 25,000 元，標準差 = 1,200 元，試求甲公司與乙公司月薪差之 98% 的信賴區間。根據此結論，我們是否可宣稱甲乙公司月薪無顯著差別？

解：(a) 本題對甲乙兩公司的分配型態及標準差均未知，但由於樣本夠大 ($n_1 = 200, n_2 = 100$ 均大於 30)，可利用中央極限定理，以 Z 分配求解。

$n_1 = 200$, $\overline{X}_1 = 26{,}000$, $\hat{S}_1 = 1{,}400$

$n_2 = 100$, $\overline{X}_2 = 25{,}000$, $\hat{S}_2 = 1{,}200$

$Z_{\left(1-\frac{\alpha}{2}\right)} = Z_{0.99} = 2.33$

$$\mu_1 - \mu_2 \sim (\overline{X}_1 - \overline{X}_2) \mp Z_{\left(1-\frac{\alpha}{2}\right)} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\Rightarrow (\overline{X}_1 - \overline{X}_2) \mp Z_{\left(1-\frac{\alpha}{2}\right)} \cdot \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}}$$ (以 $\hat{S}_1^2$, $\hat{S}_2^2$ 代替 σ_1^2 , σ_2^2)

$$= (26{,}000 - 25{,}000) \mp 2.33\sqrt{\frac{1{,}400^2}{200} + \frac{1{,}200^2}{100}}$$

$$= 1{,}000 \mp 362.46$$

$\therefore 637.51 \le \mu_1 - \mu_2 \le 1{,}362.46$ (元)

(b) 依一般常識可知，若兩公司月薪無顯著差別，則 $\mu_1 - \mu_2$ 在正負值之間 (包括零)。由於 $\mu_1 - \mu_2$ 在正值之間，不包括了相差為零的情況，所以宣稱兩公司月薪無顯著差別是不可信的。

例題 8-13

甲班使用黑板教學，今隨機抽取 $n_1 = 12$ 人，期末考統計學平均成績 $= 84$ 分，標準差 $= 4$ 分，另外乙班使用電視教學，今隨機抽取 $n_2 = 18$ 人，期末考統計學平均成績 $= 77$ 分，標準差 $= 6$ 分。設甲乙兩班統計學成績呈常態分配且其變異數相等，試求兩班統計學成績差之 99% 的信賴區間。

解： ∵兩母體爲常態分配，小樣本，且變異數相等，使用 t 分配求解。

$$n_1 = 12 , \overline{X}_1 = 84 , \hat{S}_1^2 = 16$$

$$n_2 = 18 , \overline{X}_2 = 77 , \hat{S}_2^2 = 36$$

$$f = 12 + 18 - 2 = 28 , t_{(0.995,28)} = 2.763$$

$$S_P = \sqrt{\frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{(12-1)16 + (18-1)36}{12 + 18 - 2}} = 5.30$$

$$\mu_1 - \mu_2 \sim (\overline{X}_1 - \overline{X}_2) \mp t_{(1-\frac{\alpha}{2}, f)} \cdot S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= (84 - 77) \mp 2.763(5.30)(0.373) = 7 \mp 5.46$$

$$\therefore 1.54 \le \mu_1 - \mu_2 \le 12.46 \text{ (分)}$$

例題 8-14

設甲乙兩班數學成績呈常態分配且變異數相等，今自兩班抽出部分同學其成績如下表，試求兩班數學平均分數差之 95% 的信賴區間。

甲 班	44, 56, 44, 47, 46, 53, 58, 38, 41, 30, 46, 35, 49
乙 班	39, 51, 42, 57, 41, 32, 40, 39, 55, 29, 35, 47

解： 甲班：$n_1 = 13 , \bar{X}_1 = \dfrac{\Sigma X_{1n1}}{13} = 45.15 , \hat{S}_1^2 = \dfrac{1}{13-1}\Sigma(X-45.15)^2 = 63.97$

乙班：$n_2 = 12 , \bar{X}_2 = \dfrac{\Sigma X_{2n2}}{12} = 42.25 , \hat{S}_2^2 = \dfrac{1}{12-1}\Sigma(X-42.25)^2 = 76.39$

$$S_P = \sqrt{\frac{(13-1)(63.97)+(12-1)(76.39)}{13+12-2}} = 8.36$$

$$t_{\left(1-\frac{\alpha}{2}, n_1+n_2-2\right)} = t_{0.975(23)} = 2.069$$

$$\mu_1 - \mu_2 \sim (\bar{X}_1 - \bar{X}_2) \mp t_{\left(1-\frac{\alpha}{2}, n_1+n_2-2\right)} \cdot S_P \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$= (45.15-42.25) \mp (2.069)(8.36)\sqrt{\frac{1}{13}+\frac{1}{12}}$$

$$= 2.90 \mp 6.92$$

$$\therefore -4.02 \le \mu_1 - \mu_2 \le 9.82 \text{ (分)}$$

例題 8-15

下表表示甲乙兩家電影公司所生產部分影片的放映時間，設兩家公司影片放映時間趨近於常態分配，試求兩家公司影片放映時間平均數差之 90% 的信賴區間（設兩公司影片放映時間變異數不等）。

時間 / 公司	放映時間（分）						
甲公司	110	103	94	98	87		
乙公司	82	123	97	92	118	88	175

解： 本題母體變異數不等，而樣本是獨立的小樣本，母體趨近於常態分配，可用 t 分配求解。

甲公司：$n_1 = 5$, $\overline{X}_1 = \dfrac{\Sigma X_{1n1}}{5} = 98.4$

$$\hat{S}_1^2 = \frac{1}{5-1}\Sigma(X-98.4)^2 = 76.3$$

乙公司：$n_2 = 7$, $\overline{X}_2 = \dfrac{\Sigma X_{2n2}}{7} = 110.7$

$$\hat{S}_2^2 = \frac{1}{7-1}\Sigma(X-110.7)^2 = 1{,}035.9$$

$$f=\frac{\left(\frac{\hat{S}_1^2}{n_1}+\frac{\hat{S}_2^2}{n_2}\right)^2}{\frac{\left(\frac{\hat{S}_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{\hat{S}_2^2}{n_2}\right)^2}{n_2-1}}=\frac{\left(\frac{76.3}{5}+\frac{1{,}035.9}{7}\right)^2}{\frac{\left(\frac{76.3}{5}\right)^2}{5-1}+\frac{\left(\frac{1{,}035.9}{7}\right)^2}{7-1}}$$

$$=\frac{26{,}647.30}{3{,}707.89}=7.19\doteq 7\ (\text{取較小數})$$

$$t_{\left(1-\frac{\alpha}{2},f\right)}=t_{(0.95,7)}=1.895$$

$$\mu_1-\mu_2\sim(\overline{X}_1-\overline{X}_2)\mp t_{\left(1-\frac{\alpha}{2},f\right)}\sqrt{\frac{\hat{S}_1^2}{n_1}+\frac{\hat{S}_2^2}{n_2}}$$

$$=(98.4-110.7)\mp 1.895\sqrt{\frac{76.3}{5}+\frac{1{,}035.9}{7}}=-12.30\mp 24.22$$

$$\therefore -36.52\le\mu_1-\mu_2\le 11.92\ (\text{分})$$

例題 8-16

某食品公司宣稱，吃其食品一個月可減輕體重達 4.5 公斤，今抽取 7 位女性，在食用前與食用後體重之資料如下（設體重呈常態分配）：

婦女		1	2	3	4	5	6	7
體重	食用前	58.5	60.3	61.7	69.0	64.0	62.6	56.7
	食用後	60.0	54.9	58.1	62.1	58.5	59.9	54.4

試求：

(a) 食用食品前後體重之 95% 之信賴區間。

(b) 驗證該公司宣稱是否可能是對的？

解：本題食用前與食用後之樣本，是不獨立樣本，因為屬同一人。且小樣本，體重呈常態分配，故可用 t 分配處理。

(a)

	1	2	3	4	5	6	7	Σ
體重（前）	58.5	60.3	61.7	69.0	64.0	62.6	56.7	
體重（後）	60.0	54.9	58.1	62.1	58.5	59.9	54.4	
D	−1.5	5.4	3.6	6.9	5.5	2.7	2.3	24.9
D^2	2.25	29.16	12.96	47.61	30.25	7.29	5.29	134.81

$$\overline{D} = \frac{\Sigma D}{n} = \frac{24.9}{7} = 3.56$$

$$\hat{S}_D^2 = \frac{1}{n-1}\Sigma(D-\overline{D})^2 = \frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)} = \frac{7(134.81)-(24.9)^2}{7\times 6} = 7.71$$

$$\hat{S}_D = \sqrt{7.71} = 2.78$$

$$t_{\left(1-\frac{\alpha}{2},f\right)} = t_{(0.975,6)} = 2.447$$

$$\mu_D = \mu_1 - \mu_2 \sim \overline{D} \mp t_{\left(1-\frac{\alpha}{2},f\right)} \cdot \frac{\hat{S}_D}{\sqrt{n}} = 3.56 \mp 2.447\frac{2.78}{\sqrt{7}} = 3.56 \mp 2.57$$

$$\therefore 0.99 \le \mu_D \le 6.13 \text{ (公斤)}$$

(b) $\because \mu_D$ 的信賴區間包含 4.5 公斤，$\therefore$ 該公司的宣稱可能是對的。

8-6 常態母體變異數的區間估計

- **一個母體變異數的區間估計：**

$$\frac{(n-1)\hat{S}^2}{\chi^2_{(1-\frac{\alpha}{2},\, n-1)}} \le \sigma^2 \le \frac{(n-1)\hat{S}^2}{\chi^2_{(\frac{\alpha}{2},\, n-1)}}$$

- **兩個母體變異數比的區間估計：**

$$\frac{\hat{S}_1^2}{\hat{S}_2^2} \times \frac{1}{F_{(1-\frac{\alpha}{2},\, f_1,\, f_2)}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{\hat{S}_1^2}{\hat{S}_2^2} \times F_{(1-\frac{\alpha}{2},\, f_2,\, f_1)}$$

例題 8-17

甲班國文成績呈常態分配，自甲班隨機抽取 20 名學生計算其平均數為 72 分，標準差為 4 分，試求甲班國文成績變異數 σ^2 之 98% 信賴區間。

解： $\chi^2_{(0.99,19)} = 36.19$，$\chi^2_{(0.01,19)} = 7.63$

$$\frac{(n-1)\hat{S}^2}{\chi^2_{\left(1-\frac{\alpha}{2},n-1\right)}} \le \sigma^2 \le \frac{(n-1)\hat{S}^2}{\chi^2_{\left(\frac{\alpha}{2},n-1\right)}} \Rightarrow \frac{(20-1)4^2}{36.19} \le \sigma^2 \le \frac{(20-1)4^2}{7.63}$$

$$\Rightarrow 8.40 \le \sigma^2 \le 39.84$$

例題 8-18

設甲產品之重量分配呈常態分配，今隨機抽查 $n=5$ 件，其重量為 18.3, 15.2, 18.7, 16.4, 17.8(公克)，設信賴係數為 95%，試求：

(a) 甲產品重量分配變異數之信賴區間。

(b) 甲產品重量分配標準差之信賴區間。

(c) 若已知甲產品之平均重量 $\mu=17$，試求上述 (a) 之結果。

解：(a) $n=5\,,\ \overline{X}=\dfrac{\Sigma X}{5}=\dfrac{86.4}{5}=17.28$

$$\hat{S}^2=\frac{1}{5-1}\Sigma(X-17.28)^2=\frac{8.43}{4}=2.11$$

$$\chi^2_{\left(1-\frac{\alpha}{2}\,,\,n-1\right)}=\chi^2_{(0.975\,,\,4)}=11.14$$

$$\chi^2_{\left(\frac{\alpha}{2}\,,\,n-1\right)}=\chi^2_{(0.025\,,\,4)}=0.48$$

$$\frac{(n-1)\hat{S}^2}{\chi^2_{\left(1-\frac{\alpha}{2}\,,\,n-1\right)}}\le\sigma^2\le\frac{(n-1)\hat{S}^2}{\chi^2_{\left(\frac{\alpha}{2}\,,\,n-1\right)}}$$

$$\Rightarrow\frac{(5-1)2.11}{11.14}\le\sigma^2\le\frac{(5-1)2.11}{0.48}$$

$$\Rightarrow 0.76\le\sigma^2\le 17.58$$

(b) $0.76\le\sigma^2\le 17.58$

$0.87\le\sigma\le 4.19$ （ 將上式兩邊開方 ）

(c) 若已知母體之 $\mu = 17$，則自由度爲樣本數 n 並非 $n-1$，因不必受到 $\Sigma X = n\overline{X}$ 之條件限制。

$$\therefore \text{爲}\ \frac{nS^2}{\chi^2_{(1-\frac{\alpha}{2},n)}} \le \sigma^2 \le \frac{nS^2}{\chi^2_{(\frac{\alpha}{2},n)}}$$

$$S^2 = \frac{1}{n}\Sigma(X-17)^2 = \frac{1}{5}\Sigma(X-17)^2 = \frac{8.82}{5} = 1.76$$

$$\chi^2_{(0.975,5)} = 12.83,\ \chi^2_{(0.025,5)} = 0.83$$

$$\frac{5(1.76)}{12.83} \le \sigma^2 \le \frac{5(1.76)}{0.83} \quad \Rightarrow \quad 0.69 \le \sigma^2 \le 10.60$$

[註：上式 σ^2 之信賴區間，事實上我們可以下式計算即可]

① 未知 μ：

$$\frac{\Sigma(X-\overline{X})^2}{\chi^2_{(1-\frac{\alpha}{2},n-1)}} \le \sigma^2 \le \frac{\Sigma(X-\overline{X})^2}{\chi^2_{(\frac{\alpha}{2},n-1)}} \quad [\because (n-1)\hat{S}^2 = \Sigma(X-\overline{X})^2]$$

② 已知 μ：

$$\frac{\Sigma(X-\mu)^2}{\chi^2_{(1-\frac{\alpha}{2},n)}} \le \sigma^2 \le \frac{\Sigma(X-\mu)^2}{\chi^2_{(\frac{\alpha}{2},n)}} \quad [\because n\hat{S}^2 = \Sigma(X-\mu)^2]$$

例題 8-19

自兩獨立常態母體中抽取兩組隨機樣本，第 1 組：$n_1=9, \overline{X}_1=64, \hat{S}_1^2=36$；第 2 組：$n_2=13, \overline{X}_2=59, \hat{S}_2^2=25$，試求

(a) $\dfrac{\sigma_1^2}{\sigma_2^2}$ 之 98% 的信賴區間。

(b) 是否能宣稱 σ_1^2 可能等於 σ_2^2？

(c) $\mu_1-\mu_2$ 之信賴區間。

解：(a) $f_1 = n_1 - 1 = 9 - 1 = 8 \,,\, f_2 = n_2 - 1 = 13 - 1 = 12$

$$F_{\left(1-\frac{\alpha}{2},\, f_1,\, f_2\right)} = F_{(0.99\,,\,8\,,\,12)} = 4.4994$$

$$F_{\left(1-\frac{\alpha}{2},\, f_2,\, f_1\right)} = F_{(0.99\,,\,12\,,\,8)} = 5.6668$$

$$\frac{\hat{S}_1^2}{\hat{S}_2^2} \cdot \frac{1}{F_{\left(1-\frac{\alpha}{2},\, f_1,\, f_2\right)}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{\hat{S}_1^2}{\hat{S}_2^2} \cdot F_{\left(1-\frac{\alpha}{2},\, f_2,\, f_1\right)}$$

$$\Rightarrow \frac{36}{25} \cdot \frac{1}{4.4994} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{36}{25} \cdot 56668 \Rightarrow 0.32 \le \frac{\sigma_1^2}{\sigma_2^2} \le 8.16$$

(b) ∵信賴區間在 0.32~8.16 之間，包括 $\frac{\sigma_1^2}{\sigma_2^2} = 1$，∴宣稱兩變異數相等是可能的。

(c) $\because \sigma_1^2 = \sigma_2^2 = \sigma^2$ (由 (b) 之答案)，兩獨立常態母體，且小樣本，可用 t 分配解之。

$f = 9 + 13 - 2 = 20$ ， $t_{(0.99\,,\,20)} = 2.528$

$$S_P = \sqrt{\frac{(9-1)36 + (13-1)25}{9+13-2}} = 5.422$$

$$\mu_1 - \mu_2 \sim (\overline{X}_1 - \overline{X}_2) \mp t_{\left(1-\frac{\alpha}{2}\,,\,f\right)} \cdot S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= (64 - 59) \mp (2.528)(5.422)(0.434) = 5 \mp 5.95$$

$$\therefore -0.95 \le \mu_1 - \mu_2 \le 10.95$$

8-7 點二項母體分配比例之區間估計

- 一個母體比例的區間估計：

$$P \sim \hat{P} \mp \sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

- 兩個母體比例差的區間估計：

$$(\hat{P}_1-\hat{P}_2)-Z_{(1-\frac{\alpha}{2})}\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1}+\frac{\hat{P}_2(1-\hat{P}_2)}{n_2}} \le P_1-P_2$$

$$\le (\hat{P}_1-\hat{P}_2)+Z_{(1-\frac{\alpha}{2})}\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1}+\frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}$$

例題 8-20

從高雄市人口中抽 1,000 名為樣本，發現抽煙人口有 150 名，試求高雄市抽煙人口比例之 98% 之信賴區間。

解：本題之母體爲高雄市，因不是無抽煙，就是有抽煙，爲一實驗分兩類事件，是點二項分配，$n=1{,}000\geq 30$，以 $\hat{P}$ 分配轉化爲 Z 分配求解。

$$n=1{,}000\ ,\ \hat{P}=\frac{150}{1{,}000}=0.15$$

$$Z_{(1-\frac{\alpha}{2})}=Z_{0.99}=2.325$$

$$\hat{P}-Z_{(1-\frac{\alpha}{2})}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}\leq P\leq \hat{P}+Z_{(1-\frac{\alpha}{2})}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

$$\Rightarrow\ 0.15-2.325\sqrt{\frac{0.15(1-0.15)}{1{,}000}}\leq P\leq 0.15+2.325\sqrt{\frac{0.15(1-0.15)}{1{,}000}}$$

$$\Rightarrow\ 0.15-0.026\leq P\leq 0.15+0.026$$

$$\Rightarrow\ 0.124\leq P\leq 0.176$$

例題 8-21

為估計台北市與台北縣居民贊成在附近興建核四廠的情況，台北市抽取 $n_1=100$ 人，有 52 人贊成，台北縣抽取 $n_2=125$ 人，有 34 人贊成，試求台北市與台北縣贊成比率差之 96% 的信賴區間。某專家宣稱台北市與台北縣差 20%，是否可能正確？

解：(a) $\hat{P}_1=\dfrac{52}{100}=0.52$ ， $\hat{P}_2=\dfrac{34}{125}=0.272$

$$Z_{\left(1-\frac{\alpha}{2}\right)}=z_{0.98}=2.06$$

$$P_1-P_2\sim(\hat{P}_1-\hat{P}_2)\mp Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1}+\frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}$$

$$=(0.52-0.272)\mp 2.06\sqrt{\frac{0.52\times 0.48}{100}+\frac{0.272\times 0.728}{125}}$$

$$=0.248\mp 0.132$$

$$\therefore 0.116\le P_1-P_2\le 0.38$$

(b) $\because 0.116\sim 0.38$ 包含 20%，$\therefore$某專家宣稱可能正確。

8-8 樣本大小的決定

- 估計母体平均數 μ 之樣本大小：

$$n = \left(\frac{Z_{(1-\frac{\alpha}{2})} \times \sigma}{e} \right)^2$$

- 估計母體比例 P 之樣本大小：

$$n = \frac{Z^2_{(1-\frac{\alpha}{2})} P(1-P)}{e^2} = \left[\frac{Z_{(1-\frac{\alpha}{2})} \sqrt{P(1-P)}}{e} \right]^2$$

例題 8-22

假設從一常態母體 $n(\mu,\sigma^2)$，隨機抽取 n 個樣本，計算其樣本平均數為 $\overline{X}$，試在 95% 之信賴係數下，求

(a) 若 $n=81, \sigma=9$，則以 $\overline{X}$ 估計 μ 之誤差為多少？

(b) 若欲估計誤差在 3 之內，則至少應抽出多少樣本？

(c) 設 $n=25$ 時，計算 $\hat{S}=8$，而母體標準差 σ 未知，試作上列 (2) 之結果。

解：(a) $Z_{\left(1-\frac{\alpha}{2}\right)} = Z_{0.975} = 1.96$

$$e = |\overline{X} - \mu| \leq Z_{\left(1-\frac{\alpha}{2}\right)} \cdot \frac{\sigma}{\sqrt{n}} \leq 1.96 \cdot \frac{9}{\sqrt{81}} \leq 1.96$$

(b) $n = \left[\dfrac{Z_{\left(1-\frac{\alpha}{2}\right)} \cdot \sigma}{e} \right]^2 = \left[\dfrac{1.96(9)}{3} \right]^2 = 34.57 \doteqdot 35$

(c) $n = \left[\dfrac{Z_{\left(1-\frac{\alpha}{2}\right)} \cdot \hat{S}}{e} \right]^2 = \left[\dfrac{1.96(8)}{3} \right]^2 = 27.32 \doteqdot 28$

$\therefore 28 - 25 = 3$，須再抽出 3 個樣本數。

例題 8-23

董氏基金會想了解台北市民抽煙的比例。若該基金會希望估計誤差小於 0.05 的機率為 95%，試求就下列各別情況，應取多少樣本？

(a) 根據過去資料顯示，台北市民抽煙比例為 0.2。

(b) 先抽出 35 人，計算其抽煙比例為 0.3。

(c) 台北市民抽煙比例未知。

(d) 評論 (a)(b)(c) 之結果。

解：(a) $1-\alpha=0.95$，$Z_{(1-\frac{\alpha}{2})}=Z_{0.975}=1.96$，$e=0.05$

$$P=0.2,\ n=\left[\frac{Z_{(1-\frac{\alpha}{2})}\sqrt{P(1-P)}}{e}\right]^2=\left(\frac{1.96\sqrt{0.2(0.8)}}{0.05}\right)^2=245.9\fallingdotseq 246$$

(b) $$\hat{P}=0.3,\ n=\left[\frac{Z_{(1-\frac{\alpha}{2})}\sqrt{\hat{P}(1-\hat{P})}}{e}\right]^2=\left(\frac{1.96\sqrt{0.3(0.7)}}{0.05}\right)^2=322.7\fallingdotseq 323$$

(c) P 未知，以 $P(1-P)=\frac{1}{4}$ 估計之

$$n=\left[\frac{Z_{(1-\frac{\alpha}{2})}\sqrt{P(1-P)}}{e}\right]^2=\left(\frac{1.96\sqrt{1/4}}{0.05}\right)^2=384.16\fallingdotseq 385$$

(d) 由 (a) (b) (c) 樣本大小比較，可知 (c) 之情況樣本最大 $(n=385)$。此乃以極大值 $P(1-P)=\frac{1}{4}\left(P=\frac{1}{2}\right)$ 去估計樣本 n，是最保守的作法。因此當 P 愈接近 $\frac{1}{2}$ 時，所需抽出的樣本 n 會愈大，如 (b) $P=0.3$ 比 (a) $P=0.2$，所需的樣本大 $(323>246)$。