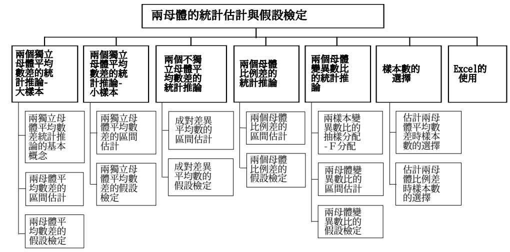
13 兩母體的統計估計與假設檢定

學習目的

1. 了解獨立大樣本及小樣本下兩母體平均數差的區間估計與假設檢定的方法、步驟及其應用。
2. 了解成對樣本成對差的區間估計與假設檢定的方法與步驟。
3. 了解兩母體比例差的區間估計與假設檢定的方法與步驟。
4. 熟悉兩樣本變異數比的抽樣分配-*F*分配。
5. 學習兩母體變異數比的假設檢定的方法、步驟及其應用。
6. 熟悉估計兩母體平均數差、比例差時樣本數的選擇。
7. 利用Excel來作兩母體差異的統計估計與假設檢定。

本章結構



兩個母體平均數差的統計推論—獨立大樣本

- **獨 立 母 體 與 不 獨 立 母 體**

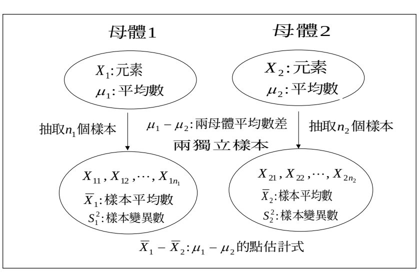
  若 X 與 Y 分別代表兩個母體的特質，若 X 與 Y 為統計獨立，則 X 與 Y 兩母體獨立，否則為不獨立。

兩個母體平均數差的統計推論—獨立大樣本

- **獨立樣本與成對樣本**

  自兩個獨立母體分別隨機抽取的樣本稱為獨立樣本。若自兩個不獨立(或成對母體)成對抽出的樣本，稱為成對樣本。

圖13.1 兩母體平均數差的估計



兩個母體平均數差的區間估計—獨立大樣本

- **$\overline{X}_1-\overline{X}_2$ 的抽樣分配**

  從兩個母體抽取兩個大且獨立的樣本，$\overline{X}_1-\overline{X}_2$ 的抽樣分配為(近似)常態分配，

  其平均數為：$\mu_{\overline{X}_1-\overline{X}_2}=\mu_1-\mu_2$

  變異數為：$\sigma^2_{\overline{X}_1-\overline{X}_2}=\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}$

- **$\sigma^2_{\overline{X}_1-\overline{X}_2}$ 的估計式**

  $\sigma^2_{\overline{X}_1-\overline{X}_2}$ 的點估計式為：

$$S^2_{\overline{X}_1-\overline{X}_2}=\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}$$

  式中：S_1^2，S_2^2 分別為樣本的變異數。

  $S_1^2=\dfrac{\sum(X_1-\overline{X}_1)^2}{n_1-1}$，$S_2^2=\dfrac{\sum(X_2-\overline{X}_2)^2}{n_2-1}$。

兩個母體平均數差的區間估計—獨立大樣本

- **獨立大樣本母體平均數差 $\mu_1-\mu_2$ 的信賴區間**

①兩母體變異數已知

$$(\bar{X}_1-\bar{X}_2)\pm Z_{\alpha/2}\sigma_{\bar{X}_1-\bar{X}_2}$$

式中：$\sigma_{\bar{X}_1-\bar{X}_2}=\sqrt{(\sigma_1^2/n_1)+(\sigma_2^2/n_2)}$

②兩母體變異數未知

$$(\bar{X}_1-\bar{X}_2)\pm Z_{\alpha/2}S_{\bar{X}_1-\bar{X}_2}$$

式中：$S_{\bar{X}_1-\bar{X}_2}=\sqrt{(S_1^2/n_1)+(S_2^2/n_2)}$

兩個母體平均數差的區間估計—獨立大樣本

- **兩樣本的混合變異數**

$$S_p^2=\frac{\sum_{i=1}^{n_1}(X_{1i}-\bar{X}_1)^2+\sum_{i=1}^{n_2}(X_{2i}-\bar{X}_2)^2}{n_1+n_2-2}=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

式中：$S_1^2=\frac{\sum_{i=1}^{n_1}(X_{1i}-\bar{X}_1)^2}{n_1-1}$，$S_2^2=\frac{\sum_{i=1}^{n_2}(X_{2i}-\bar{X}_2)^2}{n_2-1}$。

- **獨立大樣本母體平均數差 $\mu_1-\mu_2$ 的信賴區間**

兩母體變異數未知但已知相等

$$(\bar{X}_1-\bar{X}_2)\pm Z_{\alpha/2}S_{\bar{X}_1-\bar{X}_2}$$

式中：$S_{\bar{X}_1-\bar{X}_2}=S_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$　$S_p=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$

兩個母體平均數差的假設檢定—獨立大樣本

- **獨立大樣本母體平均數差 $\mu_1-\mu_2$ 的檢定統計量**

①兩母體變異數已知

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}}$$

式中：$\sigma_{\bar{X}_1-\bar{X}_2}=\sqrt{(\sigma_1^2/n_1)+(\sigma_2^2/n_2)}$

②兩母體變異數未知

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_{\bar{X}_1-\bar{X}_2}}$$

式中：$S_{\bar{X}_1-\bar{X}_2}=\sqrt{(S_1^2/n_1)+(S_2^2/n_2)}$

③兩母體變異數未知但已知相等

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_{\bar{X}_1-\bar{X}_2}}$$

式中：$S_{\bar{X}_1-\bar{X}_2}=S_p\sqrt{(1/n_1)+(1/n_2)}$

式中的 $\mu_1-\mu_2$ 為虛無假設的猜測值，S_P 為混合標準差。

圖11.2　銀行客戶等待時間差異的檢定

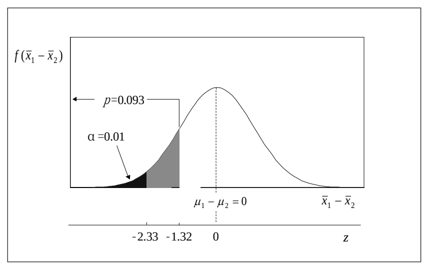


圖11.5　獨立小樣本母體平均數差的假設檢定

- 獨立小樣本
  - 母體分配已知為常態分配
    - 兩母體變異數 σ_1^2,σ_2^2 已知　$\frac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma^2}{n_2}}}\sim Z$
    - 兩母體變異數 σ_1^2,σ_2^2 未知，但相等　$\frac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t_{n_1+n_2-2}$
    - 兩母體變異數 σ_1^2，σ_2^2 未知　$\frac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{S^2}{n_1}+\frac{S^2}{n_2}}}\sim t_{\varphi}$
  - 母體分配為非常態分配
    - 兩母體變異數 σ_1^2，σ_2^2 已知　柴比氏定理
    - 兩母體變異數 σ_1^2，σ_2^2未知，但相等　無母數統計學討論
    - 兩母體變異數 σ_1^2，σ_2^2 未知　無母數統計學討論

兩個母體平均數差的統計推論—獨立小樣本

- **t分配在做母體平均數差 $\mu_1-\mu_2$ 統計推論的假設條件**

在下列假設條件為真的情況下，t分配可用來做母體平均數差 $\mu_1-\mu_2$ 的統計推論：

①兩母體為常態分配

②兩樣本為獨立小樣本($n_1<30$，$n_2<30$)

③兩個母體變異數 σ_1^2，σ_2^2 未知

兩個母體平均數差的區間估計—獨立小樣本

- **獨立小樣本常態母體平均數差的信賴區間**

①兩母體為常態且兩個變異數均已知

$(\bar{X}_1-\bar{X}_2)\pm Z_{\alpha/2}\sigma_{\bar{X}_1-\bar{X}_2}$

式中：$\sigma_{\bar{X}_1-\bar{X}_2}=\sqrt{(\sigma_1^2/n_1)+(\sigma_2^2/n_2)}$

②兩母體變異數未知

$(\bar{X}_1-\bar{X}_2)\pm t_{\varphi,\alpha/2}S_{\bar{X}_1-\bar{X}_2}$

式中：$S_{\bar{X}_1-\bar{X}_2}=\sqrt{(S_1^2/n_1)+(S_2^2/n_2)}$，$t$的自由度為$\varphi$ 。

③兩母體變異數未知但已知相等

$(\bar{X}_1-\bar{X}_2)\pm t_{v,\alpha/2}S_{\bar{X}_1-\bar{X}_2}$

式中：$S_{\bar{X}_1-\bar{X}_2}=S_p\sqrt{(1/n_1)+(1/n_2)}$，$S_P$ 為混合樣本標準差，自由度$v=n_1+n_2-2$。

兩個母體平均數差的假設檢定—獨立小樣本

- **獨立小樣本常態母體平均數差 $\mu_1-\mu_2$ 的檢定統計量**

①兩母體變異數已知

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}}$$

式中：$\sigma_{\bar{X}_1-\bar{X}_2}=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$

②兩母體變異數未知

$$t=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_{\bar{X}_1-\bar{X}_2}}$$

式中：$S_{\bar{X}_1-\bar{X}_2}=\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$ (自由度為φ)

兩個母體平均數差的假設檢定—獨立小樣本

- **獨立小樣本常態母體平均數差 $\mu_1-\mu_2$ 的檢定統計量**

③兩母體變異數未知但已知相等

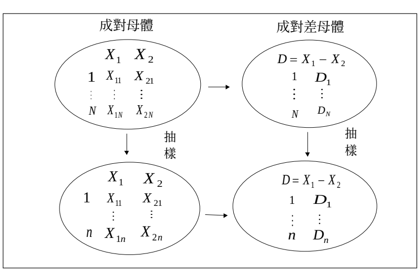
$$t=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_{\bar{X}_1-\bar{X}_2}}$$

式中：$S_{\bar{X}_1-\bar{X}_2}=S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ (自由度為$v=n_1+n_2-2$)

表13.4 兩母體平均數差異的檢定

	A	B	C	D
1	z 檢定：兩個母體平均數差異檢定			
2				
3		變數 1	變數 2	
4	平均數	1230	1050	
5	已知的變異數	90000	40000	
6	觀察值個數	25	16	
7	假設的均數差	0		
8	z	2.304684		
9	P(Z<=z) 單尾	0.010593		
10	臨界值：單尾	1.959961		
11	P(Z<=z) 雙尾	0.005296		
12	臨界值：雙尾	2.241395		

圖13.5 兩成對母體與成對樣本



兩個母體平均數差的統計推論—成對樣本

- **成對差大樣本平均數$\bar{D}$ 的平均數與變異數**

$\mu_{\bar{D}}=\mu_D$， $\sigma_{\bar{D}}^2=\frac{\sigma_D^2}{n}$

- ***t*分配的使用條件**

在下面的情況下，我們以t分配對 μ_D 進行統計估計與檢定：

①小樣本$n<30$

②母體變異數σ_D^2 未知

③成對差的母體為常態分配(即D為常態分配)

- **成對差母體變異數σ_D^2 的估計式**

$S_{\bar{D}}^2=\frac{S_D^2}{n}$

兩個母體平均數差的區間估計—成對樣本

- **成對母體平均數差 μ_D 的信賴區間**

①大樣本變異數 σ_D^2 已知

$\overline{D} \pm Z_{\alpha/2}\sigma_{\overline{D}}$ 式中：$\sigma_{\overline{D}} = \dfrac{\sigma_D}{\sqrt{n}}$

②大樣本變異數 σ_D^2 未知

$\overline{D} \pm Z_{\alpha/2}S_{\overline{D}}$ 式中：$S_{\overline{D}} = \dfrac{S_D}{\sqrt{n}}$，$S_D = \sqrt{\dfrac{\sum D^2 - (\sum D)^2/n}{n-1}}$。

③小樣本母體分配為常態分配，σ_D^2 已知

$\overline{D} \pm Z_{\alpha/2}\sigma_{\overline{D}}$

④小樣本母體分配為常態分配，σ_D^2 未知

$\overline{D} \pm t_{n-1,\alpha/2}S_{\overline{D}}$

圖11.7 成對母體平均數差的假設檢定

- 大樣本
  - σ_D^2 已知：$\dfrac{\overline{D} - \mu_D}{\frac{\sigma_D}{\sqrt{n}}} \sim Z$
  - σ_D^2 未知：$\dfrac{\overline{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} \sim Z$
- 小樣本
  - 母體常態
    - σ_D^2 已知：$\dfrac{\overline{D} - \mu_D}{\frac{\sigma_D}{\sqrt{n}}} \sim Z$
    - σ_D^2 未知：$\dfrac{\overline{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} \sim t_{n-1}$
  - 母體非常態
    - σ_D^2 已知：柴比氏定理
    - σ_D^2 未知：等級符號檢定

兩個母體平均數差的假設檢定—成對樣本

- **成對母體平均數差 μ_D 的檢定統計量**

①大樣本母體變異數已知

$$Z = \frac{\overline{D} - \mu_D}{\sigma_{\overline{D}}}$$

②大樣本母體變異數未知

$$Z = \frac{\overline{D} - \mu_D}{S_{\overline{D}}}$$

③小樣本母體常態變異數已知

$$Z = \frac{\overline{D} - \mu_D}{\sigma_{\overline{D}}}$$

兩個母體平均數差的假設檢定—成對樣本

- **成對母體平均數差 μ_D 的檢定統計量**

④小樣本母體常態變異數未知

$$t = \frac{\overline{D} - \mu_D}{S_{\overline{D}}}$$

式中：$\sigma_{\overline{D}} = \dfrac{\sigma_D}{\sqrt{n}}$，$S_{\overline{D}} = \dfrac{S_D}{\sqrt{n}}$。

⑤小樣本母體非常態，σ_p^2 已知利用柴比式定理，σ_p^2 未知，利用19章等級符號檢定。

兩個母體比例差的統計推論

- **獨立大樣本母體比例 $\hat{p}_1 - \hat{p}_2$ 的抽樣分配**

平均數 $\mu_{\hat{p}_1-\hat{p}_2} = p_1 - p_2$

變異數 $\sigma^2_{\hat{p}_1-\hat{p}_2} = \dfrac{p_1q_1}{n_1} + \dfrac{p_1q_1}{n_2}$

式中：$q_1 = 1 - p_1$，$q_2 = 1 - p_2$。

兩個母體比例差的區間估計

- **大樣本母體比例差的信賴區間**

$(\hat{p}_1 - \hat{p}_2) \pm Z_{\alpha/2}S_{\hat{p}_1-\hat{p}_2}$

式中：$S_{\hat{p}_1-\hat{p}_2} = \sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$

- **母體比例差的信賴區間**

$(\hat{p}_1 - \hat{p}_2) \pm Z_{\alpha/2}S_{\hat{p}_1-\hat{p}_2}$

式中：$S_{\hat{p}_1-\hat{p}_2} = \sqrt{\dfrac{\frac{1}{2}\times\frac{1}{2}}{n_1} + \dfrac{\frac{1}{2}\times\frac{1}{2}}{n_2}}$

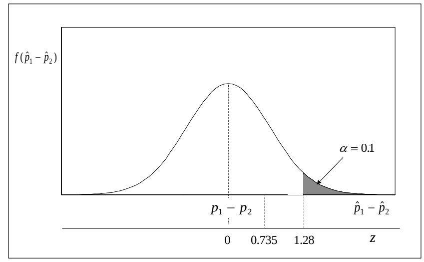
兩個母體比例差的假設檢定

- 母體比例差的檢定統計量

$$Z=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{S_{\hat{p}_1-\hat{p}_2}}$$

式中：$S_{\hat{p}_1-\hat{p}_2}=\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1}+\dfrac{\hat{p}_2\hat{q}_2}{n_2}}$

圖13.6　男女性對進口服飾偏好比例



兩個母體變異數比的統計推論

- $\dfrac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}$ 的分配

$$\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F_{n_1-1,n_2-1}$$

F 分配的性質

① F 分配為一右偏分配。F 分配決定於兩個自由度v_1，v_2，不同的v_1，v_2有不同的F分配。

② F分配的平均數與變異數

$$E(F)=\frac{v_2}{v_2-2}\quad (v_2>2),\quad V(F)=\frac{2v_2^2(v_1+v_2-2)}{v_1(v_2-2)^2(v_2-4)}\quad (v_2>4)$$

③ F 分配的定理

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F_{n_1-1,n_2-1}$$

④ $t_v^2=F_{1,v}$

意即自由度v的t隨機變數的平方恰為自由度1與v的F隨機變數。

④ F 分配的倒數性質

F的倒數仍為一F分配，其自由度為v_2,v_1，即

$$\frac{1}{F_{v_1,v_2}}\sim F_{v_2,v_1}$$

兩個母體變異數比的統計推論

- F 分配的定理

定理一

設$U\sim\chi^2_{v_1}$，$W\sim\chi^2_{v_2}$，且U與W獨立，則

$$\frac{U/v_1}{W/v_2}=\frac{\chi^2_{v_1}/v_1}{\chi^2_{v_2}/v_2}\sim F_{v_1,v_2}$$

定理二

設有兩個母體均為常態分配，即$X_1\sim N(\mu_1,\sigma_1^2)$，$X_2\sim N(\mu_2,\sigma_2^2)$且$X_1$與$X_2$互相獨立，自$X_1$，$X_2$中分別獨立隨機抽取$n_1$，$n_2$個樣本，令：

$$S_1^2=\frac{\Sigma(X_1-\overline{X}_1)^2}{n_1-1},\quad S_2^2=\frac{\Sigma(X_2-\overline{X}_2)^2}{n_2-1}$$

則

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F_{n_1-1,n_2-1}$$

圖13.7　自由度不同的*F*分配

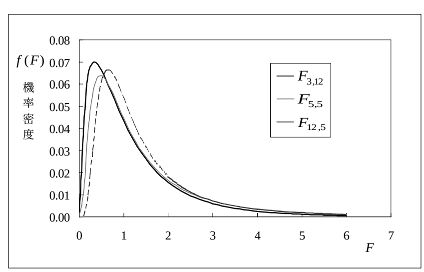
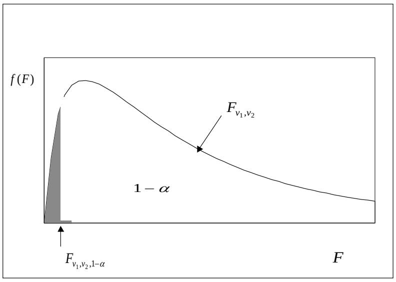


圖13.8　F分配



兩個母體變異數比的區間估計

- F 分配

$$\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n_1-1,n_2-1}$$

- 母體變異數比的信賴區間

$$\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{n_1-1,n_2-1,\alpha/2}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{n_1-1,n_2-1,1-\alpha/2}}$$

兩個母體變異數比的區間估計

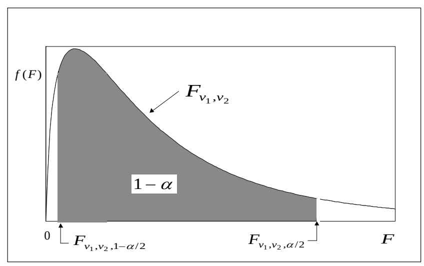
- 母體變異數比的信賴區間

$$\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{n_1-1,n_2-1,\alpha/2}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{n_1-1,n_2-1,1-\alpha/2}}$$

- F檢定統計量

$$F = \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}$$

圖13.8　$(1-\alpha)$ F的機率區間



兩個母體變異數比的統計推論

- F 檢定統計量

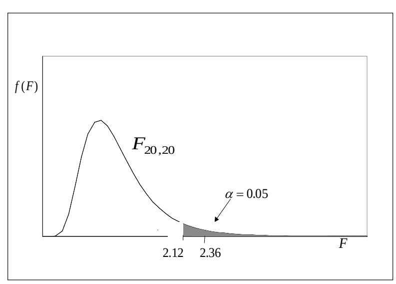
$$F = \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F_{n_1-1,n_2-1}$$

- F 檢定的決策法則

①單尾檢定：若F值 $> F_{n_1-1,n_2-1,\alpha}$，則拒絕虛無假設H_0。

②雙尾檢定：若F值 $> F_{n_1-1,n_2-1,\alpha/2}$，則拒絕虛無假設$H_0$。

圖13.12　股票投資風險的檢定



樣本數的選擇

- 估計兩母體平均數差時的樣本數

$$n \geq \frac{Z_{\alpha/2}^2(\sigma_1^2+\sigma_2^2)}{d^2}$$

- 估計兩母體比例差時樣本數

$$n \geq \frac{Z_{\alpha/2}^2(\hat{p}_1\hat{q}_1+\hat{p}_2\hat{q}_2)}{d^2} \quad 或 \quad n \geq \frac{Z_{\alpha/2}^2(2)(0.25)}{d^2}$$

表13.14　獨立小樣本母體平均數差的假設檢定

獨立小樣本	母體分配已知為常態分配	兩母體變異數σ_1^2,σ_2^2已知	$\dfrac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}} \sim Z$
		兩母體變異數σ_1^2,σ_2^2未知,但相等	$\dfrac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{S_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t_{n_1+n_2-2}$
		兩母體變異數σ_1^2, σ_2^2 未知	$\dfrac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}}} \sim t_\varphi$
	母體分配為非常態分配	兩母體變異數σ_1^2, σ_2^2 已知	柴比氏定理
		兩母體變異數σ_1^2, σ_2^2未知,但相等	無母數統計學討論
		兩母體變異數σ_1^2, σ_2^2 未知	無母數統計學討論

表13.15　成對母體平均數差的假設檢定

大樣本		σ_D^2 已知	$\dfrac{\bar{D}-\mu_D}{\sigma_{\bar{D}}} \sim Z$
		σ_D^2 未知	$\dfrac{\bar{D}-\mu_D}{S_{\bar{D}}} \sim Z$
小樣本	母體常態	σ_D^2已知	$\dfrac{\bar{D}-\mu_D}{\sigma_{\bar{D}}} \sim Z$
		σ_D^2未知	$\dfrac{\bar{D}-\mu_D}{S_{\bar{D}}} \sim t_{n-1}$
	母體非常態	σ_D^2已知	柴比 氏定理
		σ_D^2未知	等級符號檢定