第 11 章　兩母體的比較推論

Inference about Comparing Two Populations

本章比較兩個母體的四個參數：平均數差 $\mu_1 - \mu_2$（獨立樣本／成對樣本）、變異數比值 σ_1^2/σ_2^2、比例差 $p_1 - p_2$。每個推論都依「估計（信賴區間）」與「檢定（八步驟）」兩條主線。本講義依個人 # 筆記段落順序編排：先名詞速查，再核心概念、四類推論的完整推導與例題、公式總表、易錯點、計算跳板與自我檢查，供研究所備考完整複習。

目錄

1　名詞速查（五欄詞表）

全章符號先在此定義一次，後文公式不再重述。母體參數用希臘字母（μ, σ^2, p），樣本統計量用拉丁字母（$\bar{X}, S^2, \hat{P}$）。

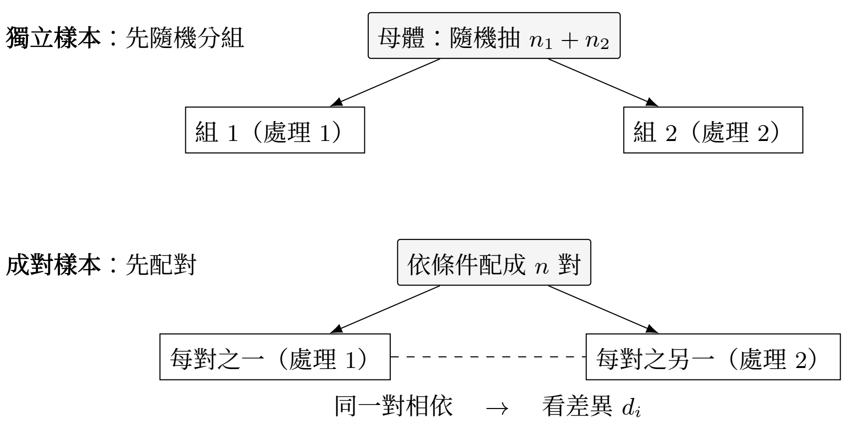
中文術語	English	符號	一句定義	用途／易混
獨立樣本	independent samples	—	兩組對象不相關，變異數可相加	對比成對樣本
成對樣本	matched pairs	$d_i = X_i - Y_i$	配對或前後測，看差異	自由度 $n-1$
平均數差	difference of means	$\mu_1 - \mu_2$	兩母體平均數之差	點估計用 $\bar{X}_1 - \bar{X}_2$
聯合變異數	pooled variance	S_p^2	等變異時兩樣本變異數的加權平均	僅 $\sigma_1^2 = \sigma_2^2$ 用
非聯合 t	unequal-variances t	ν (Satterthwaite)	變異數不等時的 t 檢定	自由度非整數
F 分配	F distribution	$F_{(\nu_1,\nu_2)}$	兩 χ^2 各除以自由度之比	右偏、查表對調
倒數性質	reciprocal property	$F_{1-\alpha} = 1/F_\alpha$	左尾臨界值取右尾倒數	自由度要對調
比例差	difference of proportions	$p_1 - p_2$	兩母體某特性比例之差	點估計用 $\hat{P}_1 - \hat{P}_2$
聯合比例	pooled proportion	$\hat{p}$	檢定 $p_1 = p_2$ 時合併估計	僅 $p_0 = 0$ 用
誤差界限	margin of error	E	信賴區間的半寬	反解樣本數

2　核心概念：兩種樣本設計與四大推論

比較兩母體前，先決定資料怎麼收集。抽樣方式不同，比較平均數時所用的抽樣分配與公式也不同，這是全章的第一個分岔點。

兩種樣本類型

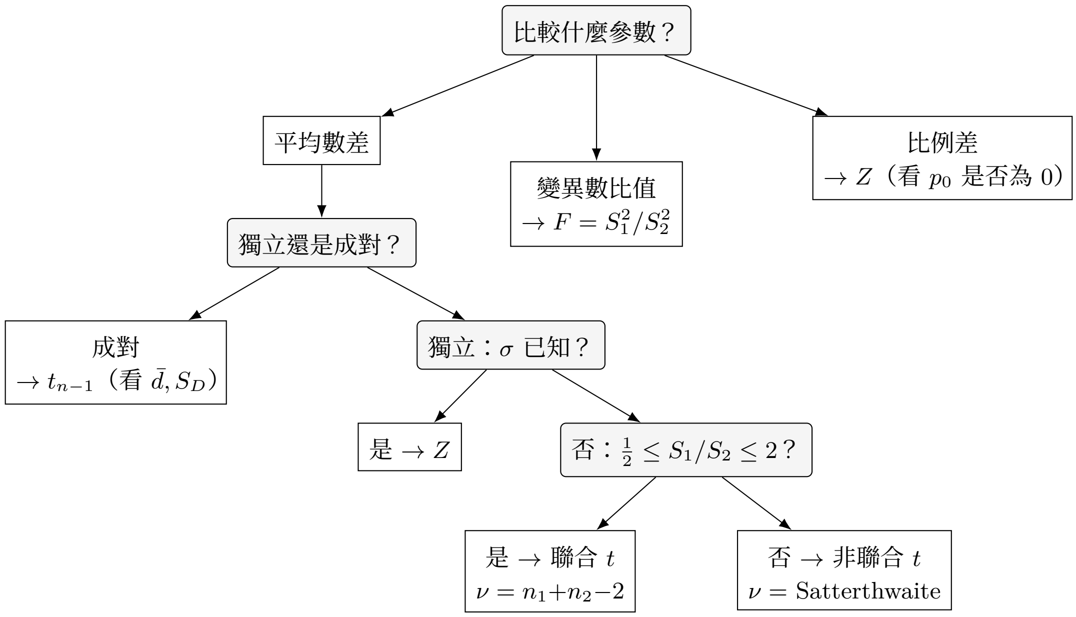
- **獨立樣本 (independent samples)**：把實驗對象隨機分成兩組，一組受處理 1、另一組受處理 2。兩組對象不同且互不相關，故兩組反應值**無關聯**，變異數可相加。例：抽兩城市的家庭各一組，比較消費差異。
- **相依（成對）樣本 (matched-pairs)**：以「成對」方式選對象，每一對在背景上相似（同性別、同年齡等），對中一個受處理 1、另一個受處理 2；或同一個體「前後測」。同一對的兩個觀察值**高度相關**，要改看其差異 d。例：以雙薪夫妻為一對，比較夫妻所得差。



本章四大推論與所在小節：

參數	樣本設計	抽樣分配／檢定
$\mu_1 - \mu_2$	獨立樣本	Z（σ 已知）／ t（σ 未知，§3）
$\mu_1 - \mu_2 = \mu_D$	成對樣本	t_{n-1}（§4）
σ_1^2/σ_2^2	兩獨立常態母體	F 分配（§5）
$p_1 - p_2$	兩獨立大樣本	Z（§6）

2.1　選擇檢定的決策流程



3　主要內容一：兩母體平均數差（獨立樣本）

以樣本平均數差 $\bar{X}_1 - \bar{X}_2$ 作為 $\mu_1 - \mu_2$ 的點估計量。

推導：$\bar{X}_1 - \bar{X}_2$ 的期望值與變異數

期望值用線性：$E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2$。
變異數：因兩組為**獨立**樣本，$\bar{X}_1, \bar{X}_2$ 互相獨立，獨立隨機變數差的變異數等於變異數**相加**：

$$V(\bar{X}_1 - \bar{X}_2) = V(\bar{X}_1) + V(\bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

(這正是「獨立」這個設計條件在公式上的體現；成對樣本不獨立，就不能這樣相加，見 §4。) 兩母體（近似）常態時 $\bar{X}_1 - \bar{X}_2$ 亦為常態。

$\bar{X}_1 - \bar{X}_2$ 的抽樣分配

$$E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2, \quad \sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \quad \bar{X}_1 - \bar{X}_2 \sim N\big(\mu_1 - \mu_2,\ \sigma_{\bar{X}_1 - \bar{X}_2}\big).$$

選哪個檢定，取決於 σ 是否已知、以及兩變異數是否相等。下面依三種情況分述。

3.1 情況一：兩母體標準差 σ_1, σ_2 已知 $\Rightarrow Z$ 檢定

標準化即得標準常態：

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \;\sim\; N(0,1).$$

推導：信賴區間怎麼來

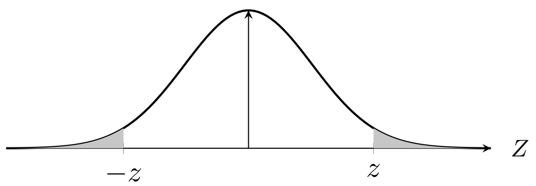
由 $P(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}) = 1-\alpha$，把 Z 代入並對 $\mu_1 - \mu_2$ 解出：

$$P\big((\bar{X}_1 - \bar{X}_2) - Z_{\alpha/2}\,\sigma_{\bar{X}_1-\bar{X}_2} \le \mu_1 - \mu_2 \le (\bar{X}_1 - \bar{X}_2) + Z_{\alpha/2}\,\sigma_{\bar{X}_1-\bar{X}_2}\big) = 1-\alpha.$$

中間夾住的就是 $\mu_1 - \mu_2$ 的 $100(1-\alpha)\%$ 信賴區間。所有「估計式 $\pm$ 臨界值 $\times$ 標準誤」的信賴區間都是這樣由樞紐量（pivot）反解而來。

$\mu_1 - \mu_2$ 的 $100(1-\alpha)\%$ 信賴區間（σ 已知）

$$(\bar{X}_1 - \bar{X}_2) \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$



雙尾檢定 $H_1: \mu_1 \ne \mu_2$ 的拒絕域 $R = \{|Z| \ge Z_{\alpha/2}\}$（兩側陰影各 $\alpha/2$）。

檢定型態	假設	檢定統計量	拒絕域 R
右尾	$H_0: \mu_1 \le \mu_2;\ H_1: \mu_1 > \mu_2$		$R = \{Z \ge Z_\alpha\}$
左尾	$H_0: \mu_1 \ge \mu_2;\ H_1: \mu_1 < \mu_2$	$Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$R = \{Z \le -Z_\alpha\}$
雙尾	$H_0: \mu_1 = \mu_2;\ H_1: \mu_1 \ne \mu_2$		$R = \{\lvert Z\rvert \ge Z_{\alpha/2}\}$

補充範例：城鄉高齡人口休閒時間（Z 檢定 + 信賴區間）

兩地區各抽 35 位 65 歲以上民眾。城市平均休閒 39.6 小時、鄉村 35.4 小時；已知 $\sigma_1 = 6.3$、$\sigma_2 = 5.8$（常態），$\alpha = 0.05$。

(1) 檢定兩地平均休閒時間是否有差異？

(1) $H_0: \mu_1 = \mu_2$ (2) $H_1: \mu_1 \ne \mu_2$

(3) $Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$ (4) $\alpha = 0.05$

(5) $R = \{|Z| \ge Z_{0.025} = 1.96\}$

(6) $z = \dfrac{39.6 - 35.4 - 0}{\sqrt{6.3^2/35 + 5.8^2/35}} = \dfrac{4.2}{1.447} = 2.9 \in R$

(7) **Reject** H_0 (8) $\alpha = 0.05$ 下有充分證據說明兩地高齡人口平均休閒時間有差異。

(2) $\mu_1 - \mu_2$ **的** 95% **信賴區間**：$4.2 \pm 1.96(1.447) = (1.4,\ 7.0)$。區間不含 0，與「拒絕 H_0」一

致。

3.2　情況二：σ 未知且 $\sigma_1^2=\sigma_2^2 \Rightarrow$ 聯合 t 檢定 (pooled t-test)

推導：為什麼要「聯合」變異數

兩母體變異數未知但相等（共同值 σ^2）時，兩個樣本變異數 S_1^2, S_2^2 都在估計同一個 σ^2。與其只用其中之一，不如把兩者以**自由度加權平均**，得到更精確的聯合估計：

$$S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}.$$

分母 n_1+n_2-2 是兩組自由度之和，也成為 t 分配的自由度。

聯合變異數與 t 統計量（$\sigma_1^2=\sigma_2^2$）

$$S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2},\qquad t=\frac{(\bar X_1-\bar X_2)-(\mu_1-\mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}\ \sim\ t_\nu,\quad \nu=n_1+n_2-2.$$

信賴區間：$(\bar X_1-\bar X_2)\pm t_{\alpha/2,\,n_1+n_2-2}\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$。

例題 11.1：直接購買 vs 委託經紀人購買基金（報酬率）

各抽 50 戶，$n_1=n_2=50$。算得 $\bar X_1=6.63,\ S_1^2=37.49$；$\bar X_2=3.72,\ S_2^2=43.34$。
先判斷變異數是否相等：$\dfrac{S_1}{S_2}=\sqrt{37.49/43.34}=0.93\in[\frac{1}{2},2]$，故視為 $\sigma_1^2=\sigma_2^2$，採**聯合 t 檢定**(雙尾)。

(1) $H_0:\mu_1=\mu_2$; (2) $H_1:\mu_1\neq\mu_2$; (4) $\alpha=0.05$

(3) $T=\dfrac{\bar X_1-\bar X_2-0}{\sqrt{S_p^2(1/n_1+1/n_2)}}$　(5) $R=\{|T|\geq t_{0.025,98}\approx 1.984\}$

(6) $S_p^2=\dfrac{49(37.49)+49(43.34)}{98}=40.42,\quad T=\dfrac{6.63-3.72}{\sqrt{40.42(1/50+1/50)}}=2.29\in R$

(7) **Reject** H_0　(8) 有充分證據說明「直接購買」的報酬率顯著高於「委託經紀人」。

95% **信賴區間**：$(6.63-3.72)\pm 1.984\sqrt{40.42(1/50+1/50)}=(0.39,\ 5.43)$。

3.3　情況三：σ 未知且 $\sigma_1^2\neq\sigma_2^2 \Rightarrow$ 非聯合 t 檢定

變異數未知且不相等時，不能合併，直接用各自的 S_i^2，自由度改用 Satterthwaite 近似 (通常非整數，取整且偏小)：

非聯合 t 統計量與自由度（$\sigma_1^2 \neq \sigma_2^2$）

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \sim t_\nu, \qquad \nu = \frac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{1}{n_1-1}\left(\dfrac{S_1^2}{n_1}\right)^2 + \dfrac{1}{n_2-1}\left(\dfrac{S_2^2}{n_2}\right)^2}.$$

信賴區間：$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2,\nu}\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}$。

例題 11.2：家族企業由子女或外部人士擔任 CEO

$n_1 = 42$（子女），$n_2 = 98$（外部）。$\bar{X}_1 = -0.1,\ S_1^2 = 3.79$；$\bar{X}_2 = 1.24,\ S_2^2 = 8.03$。變異數不相等，採**非聯合** t。

(1) $H_0 : \mu_1 = \mu_2$; (2) $H_1 : \mu_1 \neq \mu_2$; (4) $\alpha = 0.05$

(5) $\nu = \dfrac{(3.79/42 + 8.03/98)^2}{\frac{1}{41}(3.79/42)^2 + \frac{1}{97}(8.03/98)^2} \approx 111, \quad R = \{|T| \geq t_{0.025,111} \approx 1.982\}$

(6) $T = \dfrac{(-0.1 - 1.24) - 0}{\sqrt{3.79/42 + 8.03/98}} = -3.22 \in R$

(7) **Reject** H_0 (8) 有充分證據說明子女與外部人士擔任 CEO 對營業收入有顯著不同。

95% **信賴區間**：$(-0.1 - 1.24) \pm 1.982\sqrt{3.79/42 + 8.03/98} = (-2.16,\ -0.52)$。

決策法則：要不要合併變異數？

比較 $\dfrac{S_1}{S_2}$ 與 1：若 $\dfrac{1}{2} \leq \dfrac{S_1}{S_2} \leq 2$，可合理相信 $\sigma_1 = \sigma_2$，用**聯合** t；否則用**非聯合** t。更嚴謹的判斷是做 §5 的兩母體變異數比值檢定。

例題 11.4：MBA 主修財務 vs 行銷的薪資（獨立樣本，聯合 t）

各 25 人，獨立樣本。$\bar{X}_1 = 65{,}624,\ S_1^2 = 360{,}433{,}294$；$\bar{X}_2 = 60{,}423,\ S_2^2 = 262{,}228{,}559$。$\dfrac{S_1}{S_2} = 1.17 \in [\frac{1}{2}, 2]$，採聯合 t（右尾）。

(1) $H_0 : \mu_1 \leq \mu_2$; (2) $H_1 : \mu_1 > \mu_2$; (4) $\alpha = 0.05$; (5) $R = \{T \geq t_{0.05,48} = 1.676\}$

(6) $S_p^2 = \dfrac{24(360433294) + 24(262228559)}{48} = 311{,}330{,}926, \quad T = \dfrac{65624 - 60423}{\sqrt{311330926(1/25 + 1/25)}} = 1.04 \notin R$

(7) **Do Not Reject** H_0 (8) 無充分證據說明主修財務薪資高於主修行銷。

補充範例：動態 vs 靜態系統的手術時間（非聯合，含自由度計算）

動態系統 $\bar{x}_1 = 394.6,\ s_1 = 84.7,\ n_1 = 14$；靜態系統 $\bar{x}_2 = 468.3,\ s_2 = 38.2,\ n_2 = 6$（常態）。由箱型圖判斷兩母體標準差不相等，採非聯合：

$$\nu = \frac{\left(\frac{84.7^2}{14} + \frac{38.2^2}{6}\right)^2}{\frac{1}{13}\left(\frac{84.7^2}{14}\right)^2 + \frac{1}{5}\left(\frac{38.2^2}{6}\right)^2} \approx 17, \quad t_{0.025,17} = 2.11.$$

$\mu_1-\mu_2$ 的 95% 信賴區間 $=(394.6-468.3)\pm 2.11\sqrt{\frac{84.7^2}{14}+\frac{38.2^2}{6}}=(-89.4,\ -15.7)$。區間全為負，表示動態系統的手術時間顯著較短。

4 主要內容二：兩母體平均數差（成對樣本）

樣本以成對或匹配方式取得（如「前後測」或同條件配對），稱配對抽樣。

推導：為什麼把差異當成單一母體

令每對差異 $d_i=X_i-Y_i$。由於 $\mu_1-\mu_2=E(X_i)-E(Y_i)=E(d_i)=\mu_D$，比較兩母體平均數就等於檢定單一母體平均數 μ_D 是否為 0，問題回到第 8–9 章。
配對的好處：若每對 (X_i,Y_i) 高度正相關（配對成功），個體間的大變異會在相減時抵消，d_i 的變異遠小於獨立樣本，檢定更靈敏。代價是自由度只剩 $n-1$（n 為對數）。

成對差異的統計量與抽樣分配

$$d_i=X_i-Y_i,\quad \bar d=\frac{\sum_{i=1}^n d_i}{n},\quad S_D=\sqrt{\frac{\sum_{i=1}^n(d_i-\bar d)^2}{n-1}},\quad T=\frac{\bar d-\mu_D}{S_D/\sqrt n}\sim t_{n-1}.$$

μ_D 的信賴區間 $\bar d\pm t_{\alpha/2,n-1}\dfrac{S_D}{\sqrt n}$。

例題 11.5 ／ 11.6：以 GPA 配對的財務 vs 行銷薪資（成對樣本）

以 GPA 區間配對選 25 對學生，**屬相依樣本**。令 $\mu_D=\mu_1-\mu_2$，算得 $n=25,\ \bar d=5{,}065,\ S_D=6{,}647$。

1. $H_0:\mu_D=0$ 2. $H_1:\mu_D>0$ 3. $T=\dfrac{\bar d-0}{S_D/\sqrt n},\ df=24$ 4. $\alpha=0.05$
5. $R=\{T\ge t_{0.05,24}=1.711\}$ 6. $T=\dfrac{5065-0}{6647/\sqrt{25}}=3.81\in R$
7. **Reject** H_0 8. 有充分證據支持主修財務薪資高於主修行銷。

例題 11.6（續）μ_D 的 95% 信賴區間：$5065\pm 2.064\cdot\dfrac{6647}{\sqrt{25}}=(2321,\ 7809)$。

補充範例：礦物質對膽固醇的影響（前後測，完整計算）

6 位受試者，添加礦物質前 X_1、6 週後 X_2，$\alpha=0.1$（常態）。

受試者	1	2	3	4	5	6	$\sum D$	$\sum D^2$
Before X_1	210	235	208	190	172	244		
After X_2	190	170	210	188	173	228		
$D=X_1-X_2$	20	65	-2	2	-1	16	100	4890

$\bar d=\dfrac{100}{6}=16.7,\quad S_D=\sqrt{\dfrac{4890-100^2/6}{5}}=25.4$。
$H_0:\mu_D=0,\ H_1:\mu_D\ne 0$；$R=\{|T|\ge t_{0.05,5}=2.015\}$； $t=\dfrac{16.7-0}{25.4/\sqrt6}=1.61\notin R\Rightarrow$ **Do not reject** H_0：無充分證據顯示礦物質可改變膽固醇水準。

練習題（附解）：左撇子的左、右手握力

觀察 10 位左撇子的左手與右手握力，配對如下，$\alpha = 0.05$ 檢定左手握力是否大於右手。

受試者	1	2	3	4	5	6	7	8	9	10
左手	140	90	125	130	95	121	85	97	131	110
右手	138	87	110	132	96	120	86	90	129	100
$d =$ 左 $-$ 右	2	3	15	-2	-1	1	-1	7	2	10

解：成對設計。$\sum d = 36,\ \bar{d} = 3.6$；$\sum(d-\bar{d})^2 = 295.4$，$S_D = \sqrt{295.4/9} = 5.73$。
$H_0: \mu_D \le 0,\ H_1: \mu_D > 0$（右尾）；$R = \{T \ge t_{0.05,9} = 1.833\}$；$T = \dfrac{3.6-0}{5.73/\sqrt{10}} = 1.99 \in R \Rightarrow$ **Reject** H_0：有證據顯示左撇子左手握力大於右手。

5 主要內容三：兩母體變異數比值（F 檢定）

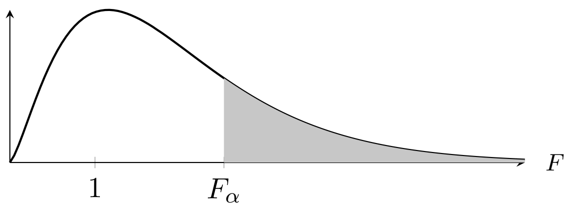
比較兩母體的「變異性」時，研究的參數是**比值** $\dfrac{\sigma_1^2}{\sigma_2^2}$，而非差 $\sigma_1^2 - \sigma_2^2$。理由：兩個變異數的差沒有已知的簡單抽樣分配，但兩個樣本變異數的**比值**有——它服從 F 分配。且 $\sigma_1^2 = \sigma_2^2$ 等價於 $\dfrac{\sigma_1^2}{\sigma_2^2} = 1$，檢定「比值是否為 1」即可。

推導：F 分配與抽樣分配

設 $\chi_1^2 \sim \chi_{\nu_1}^2,\ \chi_2^2 \sim \chi_{\nu_2}^2$ 獨立，定義 $F = \dfrac{\chi_1^2/\nu_1}{\chi_2^2/\nu_2} \sim F_{(\nu_1,\nu_2)}$。
由常態母體抽樣，$\dfrac{(n_i-1)S_i^2}{\sigma_i^2} \sim \chi_{n_i-1}^2$，代入得

$$F = \frac{\big[(n_1-1)S_1^2/\sigma_1^2\big]/(n_1-1)}{\big[(n_2-1)S_2^2/\sigma_2^2\big]/(n_2-1)} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2}{S_2^2}\cdot\frac{\sigma_2^2}{\sigma_1^2} \ \sim\ F_{(n_1-1,\,n_2-1)}.$$

特性：非負、單峰右偏、平均數近 1；有分子、分母兩個自由度。



F 分配右偏；右尾檢定的拒絕域 $R = \{F \ge F_\alpha(\nu_1,\nu_2)\}$（陰影面積 α）。圖形為示意。

倒數性質（查左尾用）與信賴區間

F 表通常只列右尾臨界值，左尾用倒數性質（兩自由度**對調**）：

$$F_{1-\alpha}(\nu_1,\nu_2) = \frac{1}{F_\alpha(\nu_2,\nu_1)}.$$

$\dfrac{\sigma_1^2}{\sigma_2^2}$ 的 $100(1-\alpha)\%$ 信賴區間：

$$\left(\frac{1}{F_{\alpha/2}(\nu_1,\nu_2)}\cdot\frac{s_1^2}{s_2^2}\ ,\quad \frac{1}{F_{1-\alpha/2}(\nu_1,\nu_2)}\cdot\frac{s_1^2}{s_2^2}\right),\quad \nu_1=n_1-1,\ \nu_2=n_2-1.$$

假設檢定的檢定統計量 $F=\dfrac{S_1^2}{S_2^2}$。

查表範例

$P(F_{(6,8)}\ge k)=0.05\Rightarrow k=F_{0.05}(6,8)=3.58$； $P(F_{(8,6)}\ge k)=0.05\Rightarrow k=F_{0.05}(8,6)=4.15$。

$k=F_{0.95}(7,10)=\dfrac{1}{F_{0.05}(10,7)}=\dfrac{1}{3.64}=0.27$。注意自由度在倒數時要對調為 $(10,7)$。

型態	假設	統計量	拒絕域 R
右尾	$H_0:\sigma_1^2\le\sigma_2^2;\ H_1:\sigma_1^2>\sigma_2^2$		$R=\{F\ge F_\alpha(n_1-1,n_2-1)\}$
左尾	$H_0:\sigma_1^2\ge\sigma_2^2;\ H_1:\sigma_1^2<\sigma_2^2$	$F=\dfrac{S_1^2}{S_2^2}$	$R=\{F\le F_{1-\alpha}(n_1-1,n_2-1)\}$
雙尾	$H_0:\sigma_1^2=\sigma_2^2;\ H_1:\sigma_1^2\ne\sigma_2^2$		$R=\{F\ge F_{\alpha/2}$ 或 $F\le F_{1-\alpha/2}\}$

例題 11.7 ／ 11.8：兩部充填機器的一致性

$n_1=n_2=25$，$s_1^2=0.6333,\ s_2^2=0.4528$。$\alpha=0.05$，檢定「第二部是否比第一部一致（變異更小）」即 $H_0:\sigma_1\le\sigma_2$ vs $H_1:\sigma_1>\sigma_2$。

(5) $R=\{F\ge F_{0.05}(24,24)=1.98\}$ (6) $F=\dfrac{0.6333}{0.4528}=1.40\notin R$

(7) **Do Not Reject** H_0 (8) 無充分證據支持第二部機器充填品質較佳。

變異數比值的 95% **信賴區間**：$\left(\dfrac{1}{2.27}\cdot\dfrac{0.6333}{0.4528},\ 2.27\cdot\dfrac{0.6333}{0.4528}\right)=(0.62,\ 3.17)$。

補充範例：兩班成績的變異數是否有差異（雙尾）

$n_1=11,\ n_2=16$，$s_1^2=182.25,\ s_2^2=457.96$，$\alpha=0.05$。

(5) $R=\{F\ge F_{0.025}(10,15)=3.06$ 或 $F\le F_{0.975}(10,15)=\frac{1}{F_{0.025}(15,10)}=\frac{1}{3.52}=0.28\}$

(6) $F=\dfrac{182.25}{457.96}=0.40\notin R$ (7) **Do Not Reject** H_0 (8) 無足夠證據證明兩班成績變異數有差異。

95% 信賴區間 $(0.13,\ 1.40)$ 包含 1，與雙尾「不拒絕 $\sigma_1^2=\sigma_2^2$」一致。

6 主要內容四：兩母體比例差（Z 檢定）

比較兩母體某特性的比例（如兩城失業率、兩族群嬰兒出生比率）。母體 1 抽 n_1、成功 X，$\hat{P}_1=X/n_1$；母體 2 抽 n_2、成功 Y，$\hat{P}_2=Y/n_2$，兩樣本獨立。

推導：$\hat{P}_1 - \hat{P}_2$ 的抽樣分配

每個樣本比例的變異數 $V(\hat{P}_i) = \dfrac{p_i(1-p_i)}{n_i}$。兩樣本獨立 ⇒ 變異數相加：

$$E(\hat{P}_1 - \hat{P}_2) = p_1 - p_2, \quad V(\hat{P}_1 - \hat{P}_2) = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}.$$

由中央極限定理，大樣本($n_i p_i \geq 5,\ n_i(1-p_i) \geq 5$)時 $\hat{P}_1 - \hat{P}_2$ 近似常態，標準化得 $Z \approx N(0,1)$。

信賴區間（一律用各自的 $\hat{P}_i$ 估計標準誤）

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

檢定的虛無假設為 $H_0: p_1 - p_2 = p_0$，依 p_0 是否為 0 分兩種檢定統計量：

推導：為什麼 $p_0 = 0$ 時要用「聯合比例」

當 $H_0: p_1 = p_2 = p$ 時，兩樣本其實來自同一個成功率 p；此時把兩樣本合併估計 p 最有效率：

$$\hat{p} = \frac{n_1\hat{P}_1 + n_2\hat{P}_2}{n_1 + n_2} = \frac{X+Y}{n_1+n_2}.$$

標準誤改用 $\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}$。這與聯合 t 檢定「合併共同變異數」是同樣的邏輯。$p_0 \neq 0$（假設 $p_1 \neq p_2$）或求信賴區間時，就**不合併**，各自用 $\hat{P}_i$。

兩種情況的檢定統計量

情況一 $p_0 = 0$（$H_0: p_1 = p_2$）： $Z = \dfrac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$, $\hat{p} = \dfrac{X+Y}{n_1+n_2}$。

情況二 $p_0 \neq 0$（$H_0: p_1 - p_2 = p_0$）： $Z = \dfrac{(\hat{P}_1 - \hat{P}_2) - p_0}{\sqrt{\dfrac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \dfrac{\hat{P}_2(1-\hat{P}_2)}{n_2}}}$。

例題 11.9：兩種沐浴香皂包裝設計的銷售比例（$p_0 = 0$）

A 超市(鮮豔設計)$n_1 = 904,\ X = 180$；B 超市(簡約綠)$n_2 = 1038,\ Y = 155$。$\hat{P}_1 = 0.1991,\ \hat{P}_2 = 0.1493$，$\hat{p} = \dfrac{180+155}{1942} = 0.1725$，$\alpha = 0.05$，右尾。

(1) $H_0: p_1 \leq p_2$; (2) $H_1: p_1 > p_2$; (5) $R = \{Z \geq Z_{0.05} = 1.645\}$

(6) $Z = \dfrac{0.1991 - 0.1493}{\sqrt{0.1725(1-0.1725)(1/904 + 1/1038)}} = 2.90 \in R$

(7) **Reject** H_0 (8) 有足夠證據說明第一種包裝銷售量顯著高於第二種。

例題 11.10 ／ 11.11（續）：要求高出超過 3%（$p_0 \neq 0$）與信賴區間

鮮豔設計成本較高，須銷售比例高出 3% 才採用。$H_0: p_1 - p_2 \leq 0.03$ vs $H_1: p_1 - p_2 > 0.03$。

$$Z = \frac{(0.1991 - 0.1493) - 0.03}{\sqrt{\frac{0.1991(1-0.1991)}{904} + \frac{0.1493(1-0.1493)}{1038}}} = 1.15 \notin R \Rightarrow \textbf{Do Not Reject } H_0.$$

例題 11.11 $p_1 - p_2$ 的 95% 信賴區間 $= (0.1991 - 0.1493) \pm 1.96\sqrt{\frac{0.1991(0.8009)}{904} + \frac{0.1493(0.8507)}{1038}} = (0.0159,\ 0.0837)$。

補充範例：候選人 A 三個月前後的支持比例（三種解法對照）

三個月前 $n_1 = 120,\ X = 55$（$\hat{P}_1 = 0.4583$）；現在 $n_2 = 80,\ Y = 41$（$\hat{P}_2 = 0.5125$），$\alpha = 0.05$。
解法一・信賴區間：$(0.4583 - 0.5125) \pm 1.96\sqrt{\frac{0.4583(0.5417)}{120} + \frac{0.5125(0.4875)}{80}} = (-0.1954,\ 0.0870)$。
解法二・以區間檢定：$0 \in (-0.1954, 0.0870) \Rightarrow$ **Do Not Reject** H_0，支持比例三個月內無顯著變化。
解法三・以統計量檢定（$p_0 = 0$）：$\hat{p} = \dfrac{55+41}{200} = 0.48$，$Z = \dfrac{0.4583 - 0.5125}{\sqrt{0.48(0.52)(1/120 + 1/80)}} = -0.75 \notin R$，三法結論一致。

推導：誤差界限與所需樣本數

要求估計誤差不超過 E、信賴水準 $1-\alpha$，由

$$E = Z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

即可反解所需樣本數（規劃調查時用，常以最保守的 $p_i = 0.5$ 估）。

7 公式整理（總表）

推論	條件	檢定統計量 / 標準誤
$\mu_1 - \mu_2$ 獨立	σ 已知	$Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$
$\mu_1 - \mu_2$ 獨立	σ 未知, $\sigma_1^2 = \sigma_2^2$	$t = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2(1/n_1 + 1/n_2)}}$, $\nu = n_1 + n_2 - 2$
$\mu_1 - \mu_2$ 獨立	σ 未知, $\sigma_1^2 \neq \sigma_2^2$	$t = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$, ν 用 Satterthwaite
μ_D 成對	差異常態	$t = \dfrac{\bar{d} - \mu_D}{S_D/\sqrt{n}}$, $\nu = n - 1$
σ_1^2/σ_2^2	兩常態母體	$F = \dfrac{S_1^2}{S_2^2}$, $(\nu_1, \nu_2) = (n_1 - 1, n_2 - 1)$
$p_1 - p_2$	大樣本, $p_0 = 0$	$Z = \dfrac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}}$
$p_1 - p_2$	大樣本, $p_0 \neq 0$	$Z = \dfrac{(\hat{P}_1 - \hat{P}_2) - p_0}{\sqrt{\hat{P}_1(1-\hat{P}_1)/n_1 + \hat{P}_2(1-\hat{P}_2)/n_2}}$

8 易錯點總整理

獨立樣本平均數差

- **選哪個檢定**：σ 已知用 Z；σ 未知先看 S_1/S_2 ——在 $[\frac{1}{2}, 2]$ 用聯合 t（$\nu = n_1 + n_2 - 2$），否則用非聯合 t（ν 用 Satterthwaite）。

- **標準誤不可混用**：聯合用 $\sqrt{S_p^2(1/n_1+1/n_2)}$，非聯合用 $\sqrt{S_1^2/n_1+S_2^2/n_2}$。
- **前提**：兩母體須（近似）常態；可畫直方圖看是否偏離鐘形。違反時改用無母數的 Wilcoxon 等級和檢定。

成對樣本

- 自由度是 $n-1$（n 是**對數**），不是 n_1+n_2-2。
- 成對 vs 獨立：題目出現「同一人前後」「配對」「匹配條件相同」即成對；分成不相關兩組才是獨立。設計不同，公式完全不同。
- 須檢查差異 d 的母體是否近似常態（畫 d 的直方圖）；違反時用 Wilcoxon 符號等級檢定。

變異數比值

- 研究對象是**比值** σ_1^2/σ_2^2，不是差。$\sigma_1^2=\sigma_2^2 \Leftrightarrow$ 比值 $=1$。
- 左尾臨界值用倒數性質 $F_{1-\alpha}(\nu_1,\nu_2)=1/F_\alpha(\nu_2,\nu_1)$，自由度要**對調**。
- 信賴區間若涵蓋 1，等價於雙尾「不拒絕 $H_0:\sigma_1^2=\sigma_2^2$」；此檢定也用來決定 §3 該用聯合或非聯合 t。

比例差

- 何時合併：只有 $H_0:p_1=p_2$（$p_0=0$）才用聯合比例 $\hat{p}$ 算標準誤；$p_0\neq 0$ 或求**信賴區間**時，一律用各自的 $\hat{P}_i$。
- 信賴區間是否含 0，等價於雙尾「$H_0:p_1=p_2$」的拒絕與否（α 對應一致時）。
- 大樣本條件：四個 $n_ip_i,\ n_i(1-p_i)$ 都要 ≥ 5，否則常態近似不可靠。

9 計算跳板（數值代入演練）

研究所計算題的關鍵在「選對公式 + 代值不出錯」。以下把每類最易考的代入步驟列成跳板，遮住右欄自行演算，再核對。

類型	已知值代入	結果
聯合變異數 S_p^2	$\dfrac{49(37.49)+49(43.34)}{50+50-2}$	$=40.42$
聯合 t 統計量	$\dfrac{6.63-3.72}{\sqrt{40.42(1/50+1/50)}}$	$=2.29$（$\geq t_{0.025,98}=1.984$，拒絕）
非聯合自由度 ν	$\dfrac{(3.79/42+8.03/98)^2}{\frac{1}{41}(3.79/42)^2+\frac{1}{97}(8.03/98)^2}$	≈ 111
成對 S_D	$\sqrt{\dfrac{4890-100^2/6}{6-1}}$	$=25.4$（$\bar{d}=100/6=16.7$）
F 統計量	$\dfrac{0.6333}{0.4528}$	$=1.40$（$< F_{0.05}(24,24)=1.98$，不拒絕）
F 左尾臨界值	$F_{0.95}(7,10)=\dfrac{1}{F_{0.05}(10,7)}=\dfrac{1}{3.64}$	$=0.27$（自由度對調）
聯合比例 $\hat{p}$	$\dfrac{180+155}{904+1038}$	$=0.1725$
比例差 Z（$p_0=0$）	$\dfrac{0.1991-0.1493}{\sqrt{0.1725(0.8275)(1/904+1/1038)}}$	$=2.90$（≥ 1.645，拒絕）

寫字區：自選一題，從假設寫到結論（八步驟）

10 自我檢查

1. 何時把兩樣本變異數合併成 S_p^2？合併後 t 的自由度是多少？
2. 成對樣本若誤用獨立樣本的公式，自由度與標準誤會出什麼錯？
3. $F_{0.95}(8,12)$ 怎麼用右尾表查？寫出倒數式。
4. 兩母體比例差的檢定，$p_0=0$ 與 $p_0=0.05$ 的標準誤算法差在哪？
5. 某 $\mu_1-\mu_2$ 的 95% 信賴區間為 $(-2.16,-0.52)$，雙尾 $\alpha=0.05$ 下能否拒絕 $H_0:\mu_1=\mu_2$？

寫字區：作答

參考答案： 1. 兩母體變異數未知且可視為相等（$\frac{1}{2} \le S_1/S_2 \le 2$）時合併，$\nu = n_1 + n_2 - 2$。 2. 自由度應為 $n-1$ 卻誤用 n_1+n_2-2（偏大），標準誤應為 $S_D/\sqrt{n}$ 卻誤用獨立樣本式，兩者都會低估顯著性風險。 3. $F_{0.95}(8,12) = 1/F_{0.05}(12,8)$，自由度對調後查右尾表。 4. $p_0 = 0$ 用聯合比例 $\hat{p}$ 算標準誤；$p_0 = 0.05$ 用各自 $\hat{P}_i$。 5. 能；區間不含 0 即拒絕 H_0。

習題索引 11-1：1, 2, 6, 7, 8 | 11-3：60, 62 | 11-4：75, 76, 78 | 11-5：86, 87, 90, 91