

Hypothesis Testing

假設檢定

Jong Yih Kuo

jykuo@ntut.edu.tw

**Department of Computer Science and
Information Engineering
National Taipei University of Technology**

基本概念

- ❑ 母群體 (population)
 - 研究所有對象組成的集合。
- ❑ 樣本 (sample)
 - 母群體 (population) 的子集合。
- ❑ 參數 (parameter, 母數)
 - 關於母體的描述性量測(descriptive measurement)。
 - 例如母體的算術平均數、標準差。
- ❑ 統計量 (statistic)
 - 樣本的描述性量測(descriptive measurement)。
 - 例如對樣本資料，算術平均數、標準差。

基本概念

❑ 變數 (variable)

- 母體 (population) 或取樣 (sample) 的某種屬性。
- 例如母體為某班級學生，研究變數為數學期末考的成績。

❑ 變數值域 (values of the variable)

- 變數所有可能呈現值的集合。
- 例如研究變數是數學期末考成績，值域為0-100的整數。

❑ 資料 (data)

- 對於一項變數，觀測到值的集合。
- 例如班級有5名學生，研究變數是數學期末考成績，取得資料為52, 76, 82, 86, 95的整數。

基本概念

□ 統計推論 (Statistical Inference)

- 根據樣本資料 (Sample)，估計、預測、決定母群體 (Population) 特性。
- 因樣本資訊小於母體資訊，推論無法完全準確。

□ 推論可靠度量度

- 信心水平 (Confidence Level)
- 顯著水準 (Significance Level)。

□ 統計推論有2種作法：

- 估計 (Estimation)
- 假設檢定 (Hypothesis Testing)。

數值描述-中央趨勢測量

❑ 算術平均數(Arithmetic Mean, Average, or Mean)

- 母體平均數 μ ，又稱期望值 $E(X)$

- $\mu = (x_1 + x_2 + \dots + x_n) / N$

- $E(X) = \mu = \sum xP(x)$

- 樣本平均數， $\bar{X} = (x_1 + x_2 + \dots + x_n) / n$

❑ 加權平均數

- $\bar{X} = (W_1x_1 + W_2x_2 + \dots + W_nx_n) / (W_1 + W_2 + \dots + W_k)$

❑ 平均數優點

- 用於統計推論，易進行代數運算，代表性強，且敏感度高。

❑ 平均數缺點

- 易受極端值影響。

數值描述-中央趨勢測量

□ 中位數 (Median)

- 給定一組資料 x_1, x_2, \dots, x_n ，由小至大排序成 $x_{(1)}, x_{(2)}, \dots, x_{(n)}$
- $me = x_{((n+1)/2)}$ ， n 是奇數。
- $me = [x_{(n/2)} + x_{(n/2+1)}]$ ， n 是偶數。

□ 中位數的特性

- 資料中有一半大於中位數，一半小於中位數。
- 優點：不易受極端值影響，適合區間資料、序位資料、無母數統計推論。
- 缺點：只考慮居中一、二個資料，缺乏敏感度。

□ 例

- 5名同學成績排序為: 50, 60, 70, 80, 90，中位數為70。
- 有1半的同學<70分；另1半的同學>70分。

數值描述-中央趨勢測量

□ 眾數(Mode)

- 資料中出現次數最多之觀測值或類別，可能不只一個。
- 優點：適用於區間資料、序位資料、類別資料。
- 缺點：一組資料的眾數是不一定，且敏感度較低。

□ 幾何平均數 (Geometric Mean)

- 適用於找出隨著時間變化的變數的成長率或改變的速率。
- 幾何平均數適用於區間資料。

數值描述-相對位置量數

▣ 百分位數 (Percentile)

- 第P百分位數(Pth percentile): 其中P%的資料<該數值，(100-P)%的資料>該數值。
- Q1，第1四分位數(first quartile,)，第25百分位數。
- Q2，第2四分位數(second quartile)，第50百分位數(50th percentile)就是中位數(median)。
- Q3，第3四分位數(third quartile)，第75百分位數，。
- 百分位數近似位置(Location of a Percentile): $L_p = (n+1)(P/100)$

▣ 四分位距 (Interquartile Range)

- 四分位距 = $Q3 - Q1$

數值描述-變異量數

□ 全距 (Range)

- Range = 資料中的最大值 - 資料中的最小值
- 優點是簡單，缺點是只考慮資料中2筆數值，包含資訊有限。

□ 變異數 (Variance)

- Population Variance: $\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$
- Sample Variance (corrected for the mean): $SS_x = S^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{n-1}$
- 算式中，取平方是因觀察距離，需避免正負數值相抵銷。
- 單位是資料單位平方，若資料單位是秒，單位是秒平方。
- 適用於比較相同型態變數的兩組以上資料。

□ 標準差 (Standard Deviation)

- Population Standard Deviation: σ
- Sample Standard Deviation: s

公式導出

$$SS_x = S^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{n-1} = \frac{1}{n-1} [\sum_{i=1}^N x_i^2 - n\bar{x}^2]$$

$$\begin{aligned} \sum_{i=1}^N (x_i - \bar{x})^2 &= \sum_{i=1}^N (x_i^2 - 2x_i\bar{x} - \bar{x}^2) \\ &= \sum_{i=1}^N x_i^2 - 2\bar{x} \sum_{i=1}^N x_i + n\bar{x}^2 \\ &= \sum_{i=1}^N x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \\ &= \sum_{i=1}^N x_i^2 - n\bar{x}^2 \end{aligned}$$

$$S_{xy} = SS_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{1}{n-1} [\sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n}]$$

$$\begin{aligned} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) &= \sum_{i=1}^n (x_i y_i - \bar{x} y_i - x_i \bar{y} - \bar{x} \bar{y}) \\ &= \sum_{i=1}^n x_i y_i - \sum_{i=1}^n \bar{x} y_i - \sum_{i=1}^n x_i \bar{y} + \sum_{i=1}^n \bar{x} \bar{y} \\ &= \sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n} - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n} + \frac{n \times \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \times n} \\ &= \sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n} \end{aligned}$$

數值描述-相對位置量數

■ 變異係數 (Coefficient of Variation)

- Population Coefficient of Variation: $CV = \sigma/\mu$

- Sample Coefficient of Variation: $cv = S/\bar{x}$

- 變異係數是一種相對差異量數，比較單位不同或單位相同但資料差異甚大的資料分散情形。

- 調查五位學生之身高及體重如下，試比較其分散程度。

- 身高：172、168、164、170、176(公分)

- 體重：62、57、58、64、64(公斤)

數值描述-線性關係量數

▣ 共變異數 (Covariance)：描述2變數關聯性，適用區間資料。

○ Population Covariance $\sigma_{xy} = \frac{\sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)}{N}$

○ Sample Covariance $S_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1}$

- 當2個變數移動方向相同，共變異數是較大正數；
- 當2個變數移動方向相反，共變異數是較大負數；
- 當2個變數關聯性較低，共變異數是較小數值。

數值描述-線性關係量數

▣ 相關係數 (Coefficient of Correlation)

○ Population Coefficient of Correlation

$$\text{➤ } \rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, -1 \leq \rho \leq +1$$

○ Sample Coefficient of Correlation

$$\text{➤ } \gamma = \frac{s_{xy}}{s_x s_y}, -1 \leq \gamma \leq +1$$

○ 相關係數接近+1，表示2變數間有正向線性關係，散佈圖 (scatter diagram) 呈現接近正斜率的直線；

○ 相關係數接近-1，表示2變數間有負向線性關係，散佈圖 (scatter diagram) 呈現接近負斜率的直線；

○ 相關係數接近0，表示2變數間沒有線性關係。

○ 有相關性不代表有因果性 (Correlation does not imply causation)。

Z分數

□ 分數落在平均數以上或以下幾個標準差的位置

○ $Z = \frac{X - \mu}{\sigma}$ 母體資料

○ $Z = \frac{X - \bar{X}}{s}$ 樣本資料

□ 特性

○ 任一組資料經Z公式轉換，均具平均數0，標準差1的特性

○ Z分數可作分配內與跨分配比較。

○ 當Z分數小於0時，表示該觀察值落在平均數以下

○ 當Z分數大於0，表示該觀察值落在平均數以上；數值越大，表示距平均數越遠，若觀察值恰等於平均數，則為0。

○ Z分數僅將原始分數進行線性轉換，未改變各分數相對關係與距離，不會改變分配形狀。

Exercise

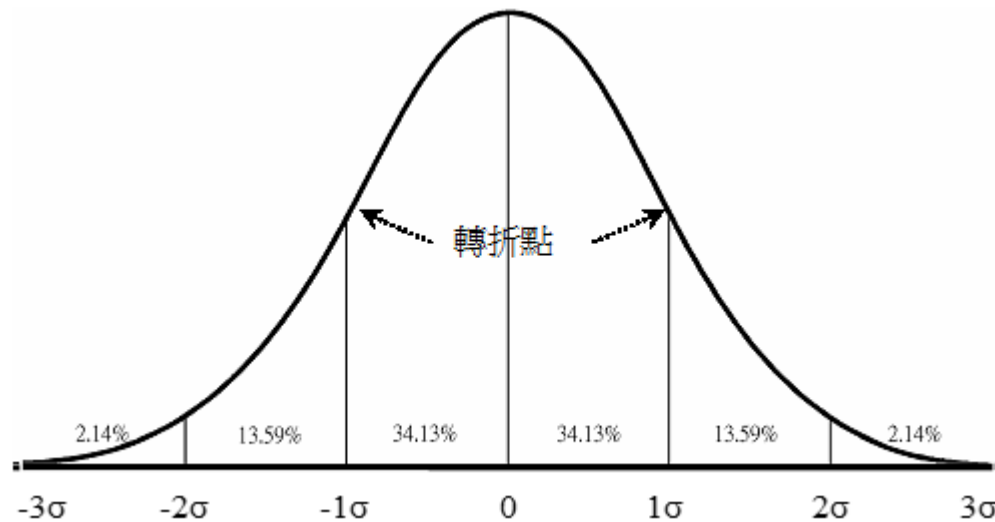
- 寫下樣本七個統計量的符號、公式與物理意義 (亦註明母體符號)

常態分佈

□ 常態分佈（Normal Distribution）

- 指一個隨機變項的觀察值，呈現對稱鐘形曲線分配
- 德國數學家 Karl F. Gauss 提出，又稱高斯分佈（Gaussian Distribution）。

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$



常態分佈

- 常態曲線沒有兩端點極限值
 - 當 $x = \mu$ 時，函數值 $f(x)$ 達最高點
 - 當 x 趨近無限大， $f(x)$ 趨近0
- 機率分配
 - 常態曲線內機率變化呈現數學規則
 - 99.7% 機率落於正負3個標準差內
 - 常態化分配全距約為6個標準差
- 轉折點（inflection points）
 - 距平均數負一個標準差位置，切線斜率由漸增轉為漸減
 - 距平均數正一個標準差位置，切線斜率由漸減轉為漸增

標準化常態分佈

■ 標準化常態分佈 (Standard Normal Distribution)

- 某變項的觀察值呈常態分配，為轉換Z分數形成的分配
- 常態分配的變數X不是原始分數，而是Z分數
- Z分數是距離平均數幾個標準差的量，不同Z值，代表距離平均值多少標準差，機率對照表可查出Z值與機率間關係

■ 標準化常態分配中

- 68.26%的觀察值落在Z值 ± 1 個標準差區間內
- 95.44%的觀察值落在Z值 ± 2 個標準差區間內
- 99.74%的觀察值落在Z值 ± 3 個標準差區間內

T分數

□ 定義

- 將Z分數以線性轉換公式轉換成平均數50，標準差10的T分數
- $T = 50 + 10Z$

□ T分數改善Z分數缺點

- Z值介於 ± 3 間，計算時帶一至二位小數點，低於平均數的Z分數有負號，實際使用較不便。

中央極限定理

□ 中央極限定理(Central Limit Theorem)

- 任一個母體，平均數 μ ，標準差 σ ，取 N 組樣本， N 夠大。
- N 夠大，取 N 組樣本，其算術平均數接近常態分佈， N 次樣本平均數的平均數為 μ ，標準差為 $\frac{\sigma}{\sqrt{n}}$ ， n 是每組樣本的個數。

□ 例，系主任稱畢業生月薪平均NT\$80,000，標準差\$10,000。

- 驗證此說法真實性，調查25位電機系畢業生，發現月薪平均為75,000。計算樣本數25，月薪平均 $\leq 75,000$ 的機率。
- X : 母體月薪。
- \bar{X} : 所有可能取樣的算術平均數，其中樣本數為25。
- 計算 $P(\bar{X} < 75,000)$
- $\mu_{\bar{X}} = \mu = 80000$ ， $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10000}{\sqrt{25}} = 2000$ 。

中央極限定理

○ 計算 $P(\bar{X} < 75,000) = P\left(\frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} < \frac{75000 - 80000}{2000}\right) = P(Z < -2.5) = 0.0062$

兩者均轉為 Z 分數

查表

○ 機率低，可認為系主任說法可能不符事實。

□ 單一母體平均數檢定-大樣本Z檢定

○ 若母體標準差 σ 已知，檢定統計量 $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

○ 若對象為大樣本($n > 30$)，且母體標準差 σ 未知，則可使用樣本標準差S替代： $Z = \frac{\bar{X} - \mu}{S/\sqrt{n}}$

○ 查『標準常態分配表』，若Z值大於查表所得臨界值，則拒絕虛無假設。

假設檢定

□ 兩種假設檢定

- rejection region method

- p-value approach

□ 統計方法進行決策過程提出兩個相反假設

- H_0 : null hypothesis (虛無假設)。

- H_1 : research or alternative hypothesis (研究假設、對立假設)。

□ 假設 H_0 為真，據此推論，可能結論有二

- 有足夠證據推論研究假設為真 → reject H_0

- reject the **null hypothesis** in favor of the alternative

- 沒有足夠證據推論研究假設為真 → not reject H_0

- not reject the **null hypothesis** in favor of the alternative

虛無假設

- 進行檢定時，將要否定(reject)之事實當作虛無假設(null hypothesis, H_0)，對立假設以 H_1 代表。
 - 檢定結果，否定虛無假設，等於接受對立假設。
 - 虛無假設與對立假設間須互斥，無重疊模糊地帶、或有無法涵蓋之真空地帶。如：
 - $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$ 。
 - 錯誤1-若安排成
 - $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \leq \mu_2$
 - 發生重疊無法互斥。
 - 錯誤2-若安排成
 - $H_0: \mu_1 < \mu_2$, $H_1: \mu_1 > \mu_2$
 - 當兩者恰好等於時，沒有被任一個假設涵蓋。

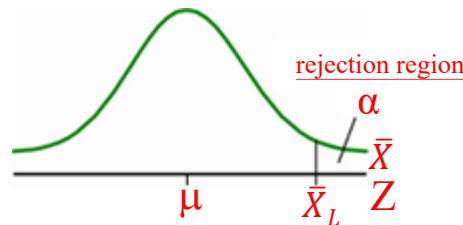
假設檢定錯誤類型

❑ 假設檢定可能的錯誤("假設 H_0 為真"的錯誤)

○ Type I error (第一型錯誤): reject a true null hypothesis.

➤ $P(\text{Type I error}) = \alpha$, α 稱為 **significance level** (顯著水準)。

➤ α 越小, Type I error 越小, 拒絕錯誤的機率越小。



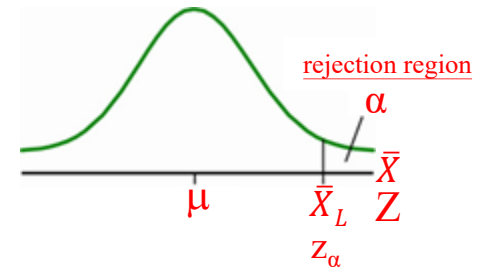
○ Type II error (第二型錯誤): don't reject a false null hypothesis.

➤ $P(\text{Type II error}) = \beta$.

		檢定結果	
		拒絕 H_0	接受 H_0
事實	H_0 為 True	錯誤拒絕 type I error	正確接受
	H_0 為 False	正確拒絕	錯誤接受 type II error

Rejection Region

- 若檢定統計量，例如樣本平均數，**落在 rejection region**
 - 推論結果，reject the null hypothesis in favor of the alternative，拒絕 H_0 ，支持 H_1 。
 - **rejection region**為 $\bar{X} > \bar{X}_L$ ， \bar{X} 為樣本平均數。
 - α 是**rejection region**面積。



- α 是 Type I error
 - 若隨機抽樣的 $\bar{X} > \bar{X}_L$ (臨界的平均值)，此時拒絕 H_0 ，犯 Type I 錯誤的機率是 α 。
 - 隨機抽出的樣本，平均值這麼極端的機率很小，此時拒絕 H_0 ，誤判的機率很小，是 α 。
 - Type I error，若 H_0 是事實，卻拒絕，這是一種錯誤。

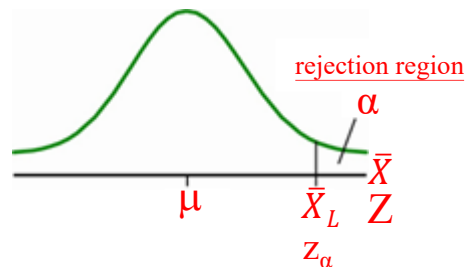
Rejection Region

□ 根據 Type I error 定義，推導出

○ $\alpha = P(\text{reject a true null hypothesis}) = P(\bar{X} > \bar{X}_L, \text{ given } H_0 \text{ is true})$
 $= P(\text{reject } H_0, \text{ given } H_0 \text{ is true}) = P(Z > z_\alpha) = \alpha$ (機率)， α 導出 z_α

○ **Z轉換**， $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ ， $z_\alpha = \frac{\bar{X}_L - \mu}{\sigma/\sqrt{n}}$

○ 計算 $P(\bar{X} > \bar{X}_L) = P(Z > z_\alpha) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{\bar{X}_L - \mu}{\sigma/\sqrt{n}}\right) = \alpha$



Rejection Region

□ 題目， $n = 400$ ， $\sigma = 65$ 。

○ H_0 : null hypothesis (虛無假設)。

➤ $\mu \leq 170$

○ H_1 : research or alternative hypothesis (研究假設、對立假設)

➤ $\mu > 170$

○ 若設定 α 為 5%，則 $z_\alpha = 1.645$ (查表，常態分布)

$$0.50 - 0.45 = 0.05 = 5\%$$

0.45 = (0.4495 + 0.4505)/2，在查表中 0.4495 ($z=1.6+0.04$) 與 0.4505 ($z=1.6+0.05$)

○ $\frac{\bar{x}_L - 170}{65/\sqrt{400}} = 1.645$, $\bar{x}_L = 175.34$ ，rejection region 為 $\bar{x} > 175.34$

○ 取得樣本算術平均數是 178，落在 rejection region，則 reject the null hypothesis，推論 alternative hypothesis $\mu > 170$ 為真。

單/雙尾檢定

□ 等於與不等於，則雙尾檢定

○ $H_0: \bar{X} = \mu$, $H_1: \bar{X} \neq \mu$

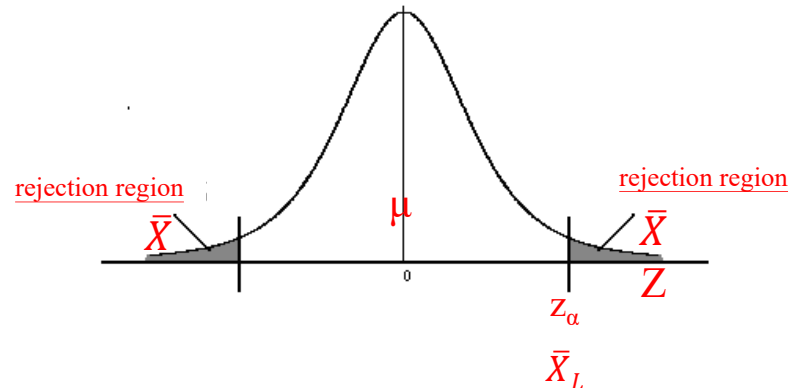
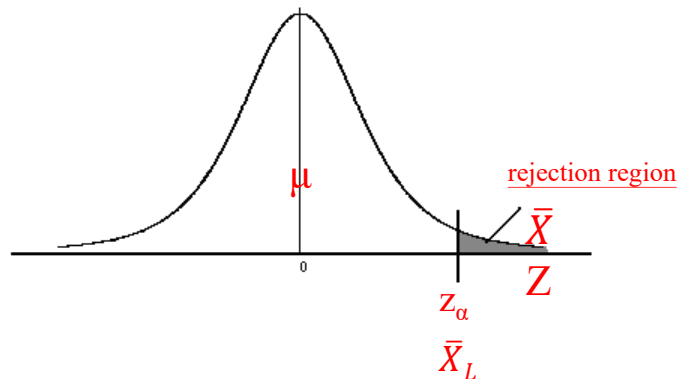
○ 無論檢定統計量觀察值 Z 落在左側或右側之 rejection region，均表示 $\bar{X} \neq \mu$ 。

➢ 落在左側，表示 $\bar{X} < \mu$ ；落在右側，表示 $\bar{X} > \mu$ 。

□ 等於與大於，則右側單尾檢定

○ $H_0: \bar{X} \leq \mu$, $H_1: \bar{X} > \mu$ 或 $H_0: \bar{X} = \mu$, $H_1: \bar{X} > \mu$

○ 當檢定統計量之觀察值 Z 落右側之 拒絕域，均表示 $\mu_1 > \mu_2$ 。

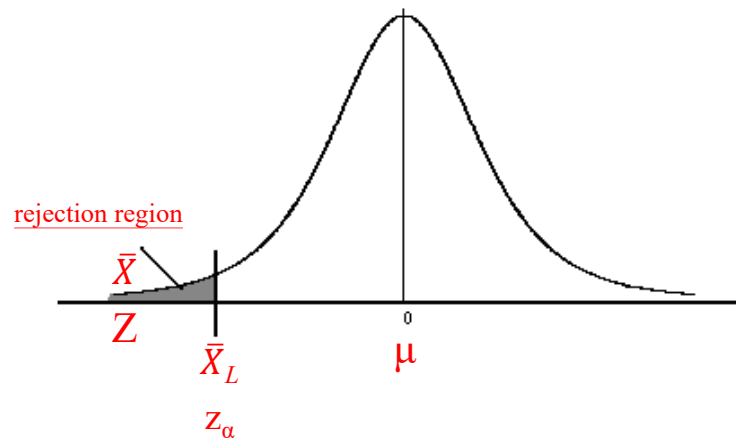


單/雙尾檢定

□ 等於與小於，則左側單尾檢定

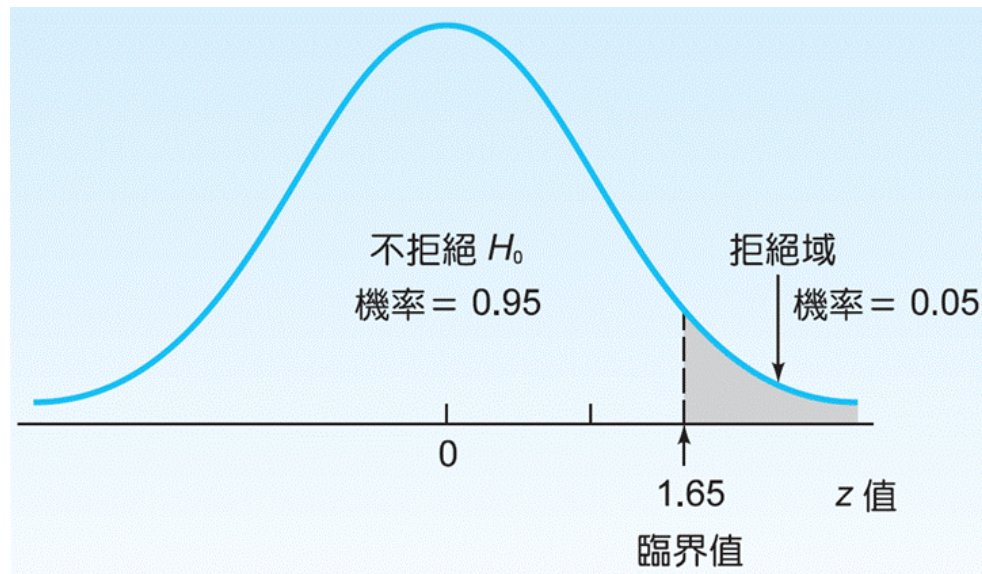
○ $H_0 : \bar{X} \geq \mu$, $H_1 : \bar{X} < \mu$ 或 $H_0 : \bar{X} = \mu$, $H_1 : \bar{X} < \mu$

○ 當檢定統計量之觀察值 Z 落左側之拒絕域，均表示 $\mu_1 < \mu_2$ 。



單尾檢定

- 當對立假設 H_1 ，表示一個方向時，即為單尾檢定，例如：
 - H_1 ：全職員工的年終獎金超過 \$35,000。 $(\mu > \$35,000)$
 - H_1 ：行駛在 I-95 公路上的卡車時數每小時小於 60 英里。 $(\mu < 60)$
 - H_1 ：對於加油的顧客，付現的人數少於 20 %。 $(\mu < 20)$
- 在 0.05 顯著水準與右尾檢定下的 z 統計量抽樣分配



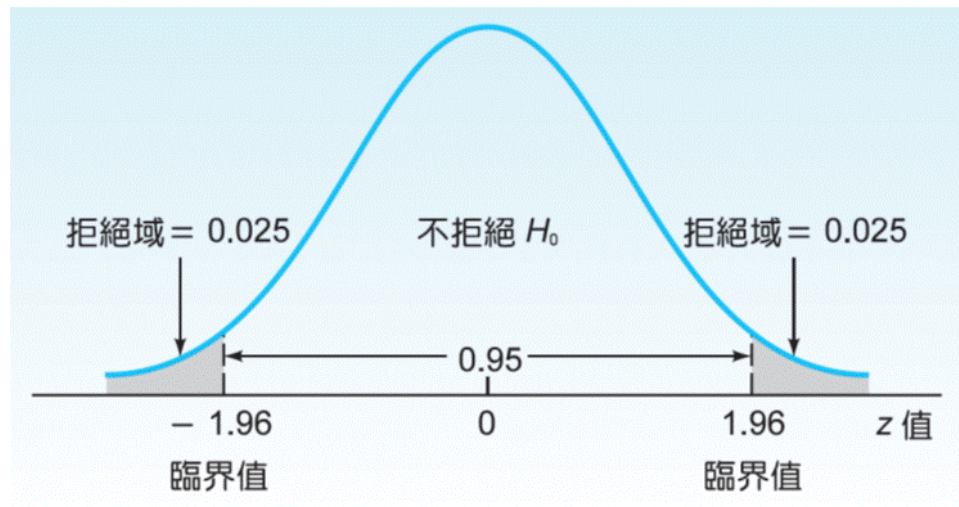
雙尾檢定

□ 如果對立假設並沒有指出一個方向時，使用雙尾檢定。例：

○ H_1 ：Georgetown 城沃爾瑪百貨每位顧客的平均消費金額不等於 \$25。 ($\mu \neq \25)

○ H_1 ：每加侖汽車售價不等於 \$1.54。 ($\mu \neq \1.54)

○ 雙尾檢定的不拒絕域與拒絕域，顯著水準為 0.05



假設檢定－檢定步驟

- ❑ 設定虛無假設 H_0
- ❑ 設定對立假設 H_1
- ❑ 決定顯著水準 (α)
- ❑ 選擇適當檢定統計量(Z, t, F)，由 α (導出)決定rejection region臨界點(critical value) z_α 。
- ❑ 計算所選之檢定統計量的觀察值(Z)
- ❑ 取得樣本，計算，查表。
- ❑ 結論：當檢定統計量的觀察值落入拒絕域，放棄虛無假設 H_0 ；反之，無法放棄虛無假設 H_0 （接受虛無假設）

假設檢定 – Type II Error 機率

□ 根據 Type II error 的定義，前例可推導出

- 若顧客平均每次消費金額(μ)高於\$180，發行NFC卡獲利會很高，使經理不願犯 Type II error，
 - $H_0: \mu \leq 180$ ， $H_1: \mu > 180$
- $\beta = P(\bar{X} < 175.34, \text{ given that the null hypothesis is false})$
- 因此 $\beta = P(\bar{X} < 175.34, \text{ given that } \mu > 180)$
- $\beta = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \frac{175.34 - 180}{65/\sqrt{400}}\right) = P(Z < -1.43) = 0.0764$
- 母群體算術平均數是180，錯誤地 not reject the null hypothesis 的機率是 0.0764

		檢定結果	
		拒絕 H_0	接受 H_0
事實	H_0 為 True	錯誤拒絕 type I error	正確接受
	H_0 為 False	正確拒絕	錯誤接受 type II error

假設檢定 – Type I II Error

- ❑ 犯下 Type I error 的機率和 Type II error 的機率關聯
 - 若試著降低犯下 Type I error 的機率 (α)，犯下 Type II error 的機率 (β) 將升高。
 - 如何取捨犯下 Type I error 及 Type II error 的機率，取決於犯下 Type I error 及 Type II error 導致的代價。
- ❑ 樣本數的影響
 - 若增加樣本數(sample size)，可在不改變 α 的情況下，降低 β 。
 - 樣本數愈大，代表資訊愈完整，犯錯機率會降低，作出判斷品質會提高。

單一母體平均數檢定-大樣本Z檢定

- 單一母體，若母體標準差 σ 已知，其各項檢定所使用之檢定統計量為：

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

- 若處理對象為大樣本（ $n > 30$ ），且母體標準差 σ 未知，則可使用樣本標準差 S 來替代：

$$Z = \frac{\bar{X} - \mu}{S / \sqrt{n}}$$

- 查『標準常態分配表』，若 Z 值大於查表所得臨界值，則放棄虛無假設。

假設檢定－範例一

- 自全校隨機抽取40位學生成績，於 $\alpha=0.05$ 之顯著水準，是否可接受全校平均成績為70分之假設。 **$H_0=70$** ， **$H_1: \mu \neq 70$**

60	80	78	70	80	80	54	78
85	88	85	85	80	85	56	82
25	85	78	75	78	82	47	83
60	80	78	70	82	78	60	75
80	88	78	83	90	90	49	82

- 計算檢定統計量的觀察值Z

$$P(\bar{X} > \bar{X}_L) = P(Z > z_\alpha) \quad P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{\bar{X}_L-\mu}{\sigma/\sqrt{n}}\right) = \alpha$$

- 將樣本數 **$n=40$** ，樣本平均數 $\bar{X}=75.5$ 與樣本標準差 **$S=13.69$** 及 $\mu=70$ ，代入檢定統計量 **Z** 公式， **$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = 2.54$**

- 結論 ($\alpha = 0.05$ ，雙尾檢定 $(1-0.05)/2 = 0.475$ ， $Z = 1.96$)

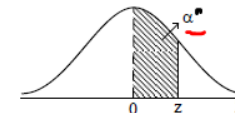
- 檢定統計量觀察值 **$Z=2.54 > 1.96$** (臨界值 **z_α**)，落入拒絕域
- 放棄虛無假設 **$H_0=70$** ，接受對立假設 **$H_1: \mu \neq 70$** 。
- 無法接受全班平均成績為70分之假設。

假設檢定－範例一

- ❑ 設定虛無假設， $H_0: \mu=70$ ， $H_1: \mu \neq 70$ ，為雙尾檢定，決定顯著水準 $\alpha=0.05$
- ❑ 選擇檢定統計量 Z ，決定拒絕域，採雙尾檢定
 - 查 $\alpha=(0.05/2)=0.025$ 。累積機率為 $0.5-0.025=0.475$ ，拒絕域臨界值 z_α 為1.96 (1.9+0.06)。
 - 檢定統計量 Z 若 <-1.96 或 >1.96 ，應放棄虛無假設。

附表1 z 分配表

$$P(0 < Z < z) = \alpha^*$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767

假設檢定－範例二

□ 已知母體標準差，檢定母體平均數 μ

- 根據財務分析，若顧客平均每次消費金額 μ 高於\$170，發行NFC卡可獲利。
- 假設每次消費金額是常態分佈，標準差\$65。取樣400人，樣本平均數(sample mean)\$178。設 α 為5%，能否推論發行NFC卡可獲利？

□ 假設檢定

$$P(\bar{X} > \bar{X}_L) = P(Z > z_\alpha) P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{\bar{X}_L - \mu}{\sigma/\sqrt{n}}\right) = \alpha$$

- $H_0: \mu = 170$ (可直接以母群體平均數為 μ 計算)
- $H_1: \mu > 170$ (單尾檢定， $0.5 - 0.05 = 0.45$ ， $z_\alpha = 1.645$ (0.450查表))
- $\frac{\bar{X}_L - 170}{65/\sqrt{400}} = 1.645$, $\bar{X}_L = 175.34$ ，rejection region為 $\bar{X} > 175.34$
- 樣本平均數是178，落在rejection region，則reject the null hypothesis，推論alternative hypothesis $\mu > 170$ 為真。

假設檢定－研究假設

□ 研究假設的方式

- 前例，若「決定發行NFC卡但實際上無法獲利」的代價(若發行NFC卡但無法獲利會賠1個資本額)比「決定不發行NFC卡但實際上可獲利」嚴重，因要避免犯下代價較高的錯誤，把目標設定為證實發NFC卡可獲利，因此假設安排如下：
 - $H_0: \mu = 170$
 - $H_1: \mu > 170$
- 若「決定不發行NFC卡但實際上可以獲利」的代價(若不發行NFC卡會少賺1個資本額)較嚴重，會把目標設定為證實發NFC卡無法獲利，假設安排如下：
 - $H_0: \mu = 170$
 - $H_1: \mu < 170$

假設檢定－範例三

- 紐約州西部 Jamestown 鋼鐵公司，主要製造及組裝書桌或其他辦公室設備。Fredonia 分廠，每週生產的 A325 型號書桌數量服從常態分配，其平均數為 200 張、標準差為 16 張書桌。最近由於市場需求增加，該工廠引進新的生產方式並雇用新員工。公司副總裁想調查 A325 型號書桌在引進新的生產方式後，產量是否有改變。在 0.01 的顯著水準下，Fredonia 分廠的平均書桌產量是否為 200 張？



假設檢定－範例三

- 步驟 1：建立虛無假設與對立假設。

- $H_0: \mu = 200$; $H_1: \mu \neq 200$

$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

- 步驟 2：選擇顯著水準，使用 0.01 的顯著水準。

- 步驟 3：選擇樣本統計量。

- 母體服從常態分配，且母體標準差 σ 已知，使用 z 分配做為檢定統計量。

- 步驟 4：決定決策法則。

- 找出 z 臨界值。此處是雙尾檢定，兩尾內的機率是 0.01 的一半 - 0.005。左右兩尾間面積 0.99 為不拒絕 H_0 的機率。曲線下右半邊面積，0.5；由 $0.5 - 0.005 = 0.495$ ，得 0 至臨界值間的面積是 0.495；查表得最接近 0.495 的機率值是 0.4951，而其相對應的 z 值為 2.58。

假設檢定－範例三

□ 顯著水準為 0.01 的決策法則

○ 若 $z > 2.58$ 或 $z < -2.58$ ，則拒絕 H_0

□ 步驟 5：做決策。 $\bar{X}=203.5, n=50$

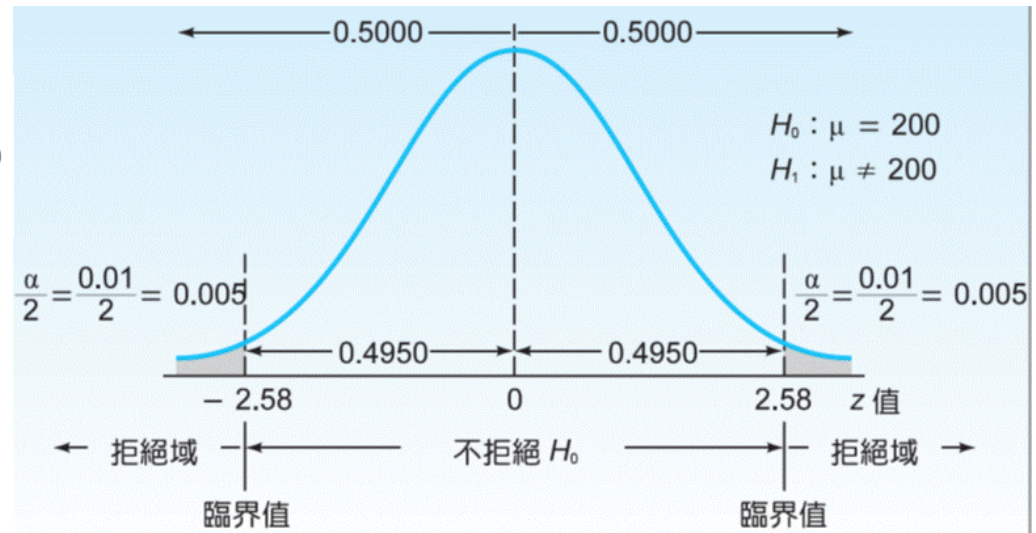
$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{203.5 - 200}{16 / \sqrt{50}} = 1.55$$

$$|Z| > Z_{\alpha/2}$$

$$\left| \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \right| > Z_{\alpha/2}$$

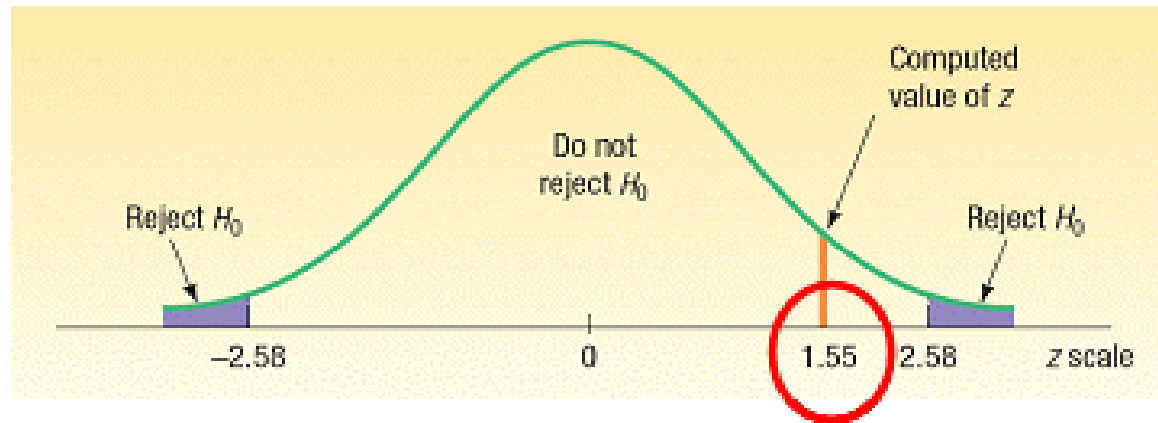
$$\left| \frac{203.5 - 200}{16 / \sqrt{50}} \right| > Z_{.01/2}$$

1.55 is not > 2.58



假設檢定－範例三

- 計算結果是 1.55 未落在拒絕域內，因此不拒絕 H_0 。結論是母體平均數與 200 張書桌沒有差異，因此根據樣本證據顯示：Fredonia 分廠的每週產量與原先的平均產量 200 張書桌沒有顯著的不同，而母體平均週產量與去年週產量間的差距 3.5 單位，可能導因於抽樣誤差。



假設檢定 – p -值法

□ 特性

- 計算樣本統計量的值至少與實際觀測值一樣大的機率（**假設虛無假設為真**），其檢定程序是將機率值，稱為 p 值，與顯著水準進行比較。
- 如果 p -值比顯著水準小，則拒絕 H_0 ；如果 p -值比顯著水準大，則不拒絕 H_0 。
- 單尾檢定： p -值 = $P\{z \geq \text{檢定統計量}\}$
- 雙尾檢定： p -值 = $P\{z \geq \text{檢定統計量的絕對值}\}$
- 假若 $z = 1.44$ ，顯著水準為 0.05，且為雙尾檢定，所以 p -值 = $P\{z \geq 1.44\} = 2(0.5 - 0.4251) = 0.1498$ 。
- 因為 $0.1498 > 0.05$ ，所以不拒絕 H_0 。

假設檢定 – p -值法

□ 特性

- 統計分析者計算統計檢定後，提供 p -value，將拒絕與不拒絕 H_0 留給決策者自行決定。
- 假設 H_0 為真，以現有樣本資料作檢定統計量，設定為臨界值的顯著水準 α ，計算出的機率。
 - 觀察到大於檢定統計量 (test statistic) 的機率。
 - 能拒絕 H_0 的最小顯著水準。
 - 可能犯 type I error 的機率。
- p -value越小，表示抽樣樣本越不可能出現，表示拒絕 H_0 不太可能錯，因此拒絕 H_0 。
 - 假設 null hypothesis 為 true 的前提，觀察到檢定統計量 (test statistic) 比取樣得到的值更極端的機率。

假設檢定 – p -值法

○ $p\text{-value} = P(\bar{X} > 178) = P\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} > \frac{178 - 170}{65/\sqrt{400}}\right) = (Z > 2.46) = 1 - p(Z < 2.16) = 1 - 0.9931 = 0.0069$

- 根據取樣分佈，當母群體算術平均數(population mean)為170時，觀察到樣本算術平均數大於178的機率是0.0069
- 因這機率很低，懷疑假設 null hypothesis 為 true 的前提，因此 reject the null hypothesis，而推論 alternative hypothesis 為真。

假設檢定 – p -值法判定標準

□ 判定標準

- < 0.01 : 有壓倒證據推論 alternative hypothesis 為真，此檢定具高顯著意義。
- $0.01 < p\text{-value} < 0.05$: 有強烈證據推論 alternative hypothesis 為真，此檢定具顯著意義。
- $0.05 < p\text{-value} < 0.10$: 有弱證據推論 alternative hypothesis 為真，此檢定未具統計意義。
- $0.10 < p\text{-value}$: 沒有證據推論 alternative hypothesis 為真，此檢定未具統計意義。
- 當母群體平均數為 170，平均值大於 178 的機率是 0.0069
 - 因這機率很低，沒有足夠證據證明 H_0 成立，拒絕 "null hypothesis 為 true"，而推論 alternative hypothesis 為真。

假設檢定－範例四

□ 系主任稱畢業生月薪平均NT\$80,000，標準差\$10,000。

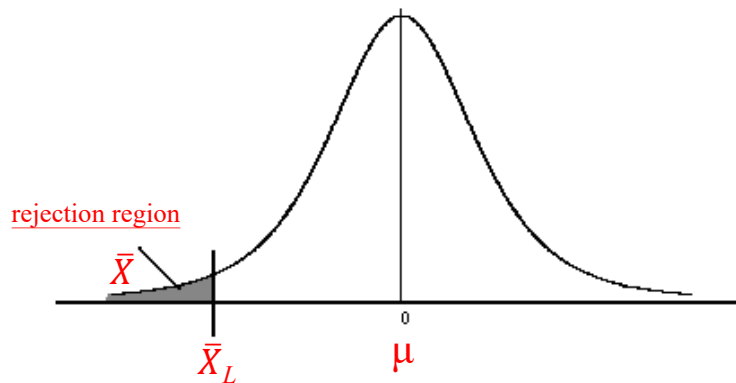
○ 為驗證此說法真實性

➤ 隨機一次調查25位電機系畢業生，發現月薪平均為75,000。

➤ 樣本數25。

➤ 計算任取一次樣本，月薪平均 $\leq 75,000$ 的機率，若機率太低表示有可能系主任所稱與事實不符。

➤ 根據中央極限定理，平均值應為常態分佈，亦即99.74%平均值為 $80,000(\mu) \pm 3$ 標準差 $(\frac{\sigma}{\sqrt{n}})$ 。



假設檢定－範例四

□ 系主任稱畢業生月薪平均NT\$80,000，標準差\$10,000。

○ X : 母體月薪。

○ \bar{X}_L : 隨機調查月薪平均 $\leq 75,000$ 。

○ \bar{X} : 任何可能取樣的月薪平均數，樣本數為25。

○ 計算 $P(\bar{X} < 75,000)$

➤ 根據中央極限定理， $\mu_{\bar{X}} = \mu = 80000$ ， $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10000}{\sqrt{25}} = 2000$ 。

○ 轉成 Z 分數比較，變成標準常態分佈，可以查標準常態分佈 Z 分數機率分配表。

○ 計算 $P(Z < \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}}) = P(Z < \frac{75000 - 80000}{2000}) = P(Z < -2.5)$

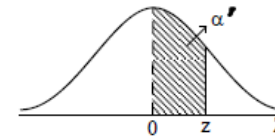
○ 機率低，可認為系主任說法可能不符事實。

假設檢定－範例四

- 計算 $P(Z < -2.5) = 0.0062$ [單尾檢定]
- $[0.5 - 0.4938 = 0.0062]$ ，機率低，可認為系主任說法可能不符事實。

附表1 z 分配表

$$P(0 < Z < z) = \alpha'$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952

假設檢定－範例五

- McFarland保險公司處理一件理賠案件需花費 \$60。與其他保險公司比較，McFarland 保險公司處理理賠案件的成本比其他公司高，因此該公司進行降低成本的計畫。為評估成效隨機抽選最近的 26 件理賠案件，其樣本資料如下：

\$45	\$49	\$62	\$40	\$43	\$61
48	53	67	63	78	64
48	54	51	56	63	69
58	51	58	59	56	57
38	76				

- 在 0.01 的顯著水平下，McFarland 保險公司的平均成本低於 \$60 嗎？

假設檢定－範例五

- 步驟 1：建立虛無假設與對立假設。
 - $H_0: \mu \geq 60$ ； $H_1: \mu < 60$
- 步驟 2：選擇顯著水準為 0.01
- 步驟 3：計算檢定統計量。t 分配，因樣本小於 30。
 - 合理假設每件理賠案件平均成本服從常態分配。
- 步驟 4：陳述決策法則。
 - Df(自由度)即樣本內觀測資料數量減去樣本的組數， $n-1$ 。
 - 樣本數量 26，組數 1，自由度 $26-1=25$ 。表中找標示單尾，顯著水準 0.01 那行，左邊找自由度 25 那列，交叉點 2.485，計算檢定統計量值。本例單尾檢定且拒絕域在左邊，臨界值負數 (-2.485)。決策法則是若 t 值小於 -2.485 須拒絕 H_0 。

假設檢定－範例五

▣ 步驟 5：做決策。

○ 26 個樣本的平均成本是 \$56.42，樣本標準差 \$10.04。

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{\$56.42 - \$60}{\$10.04/\sqrt{26}} = -1.818$$

○ 因 -1.818 落在臨界值 -2.485 的右邊，所以在 0.01 顯著水準下，不能拒絕虛無假設。表示降低成本計畫，並沒有將平均處理成本降至 \$60 以下。也就是樣本平均數 \$56.42 與母體平均數 \$60 間的差距 \$3.58，可能源自抽樣誤差。

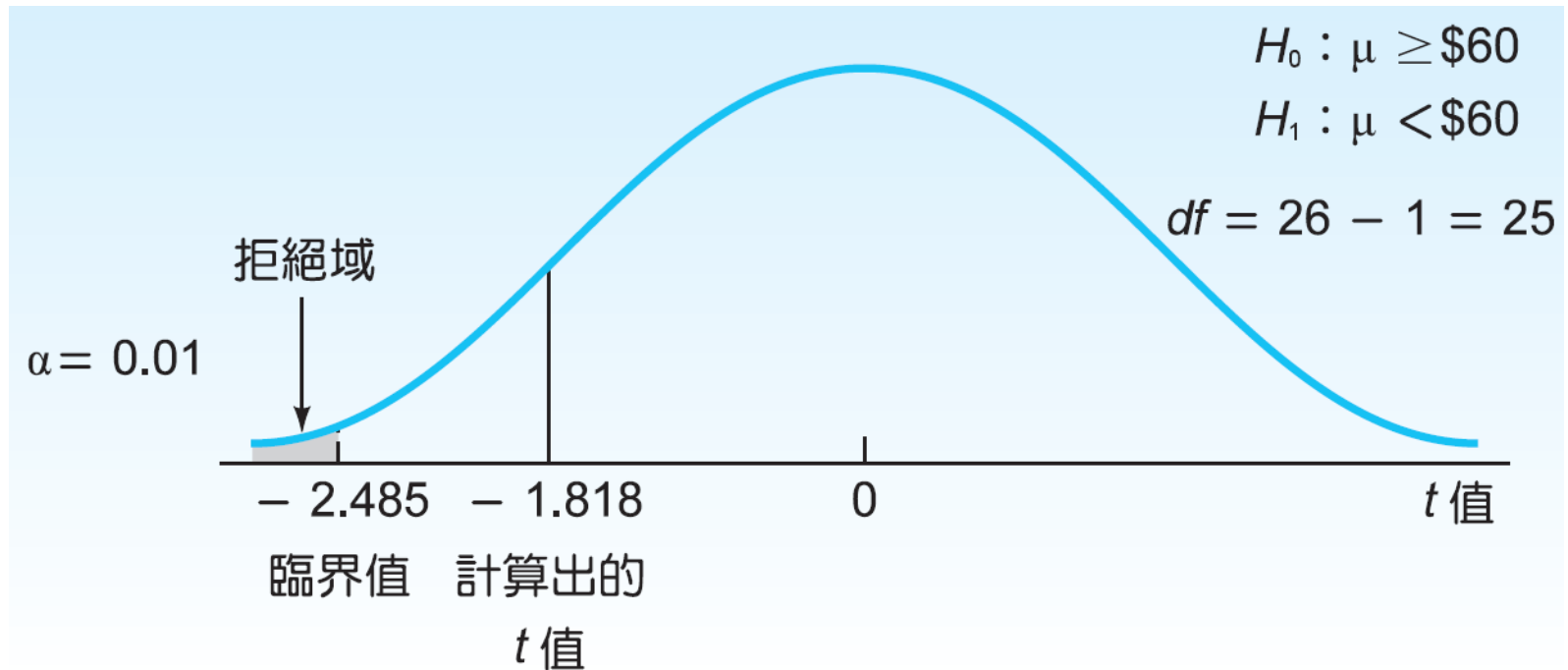
假設檢定－範例五

TABLE 10-1 A Portion of the t Distribution Table

Confidence Intervals						
	80%	90%	95%	98%	99%	99.9%
df	Level of Significance for One-Tailed Test, α					
	0.100	0.050	0.025	0.010	0.005	0.0005
	Level of Significance for Two-Tailed Test, α					
	0.20	0.10	0.05	0.02	0.01	0.001
∴	∴	∴	∴	∴	∴	∴
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.768
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646

假設檢定－範例五

□ t 分配的拒絕域，顯著水準為 0.01



假設檢定－範例六

- 一個小型平衡錘的平均長度是 43 公釐。生產主管認為調整後的機器所生產之平衡錘的長度已經改變，他要求工程部門進行調查。工程部門選取 12 個平衡錘的樣本進行測量。其測量結果如下表：

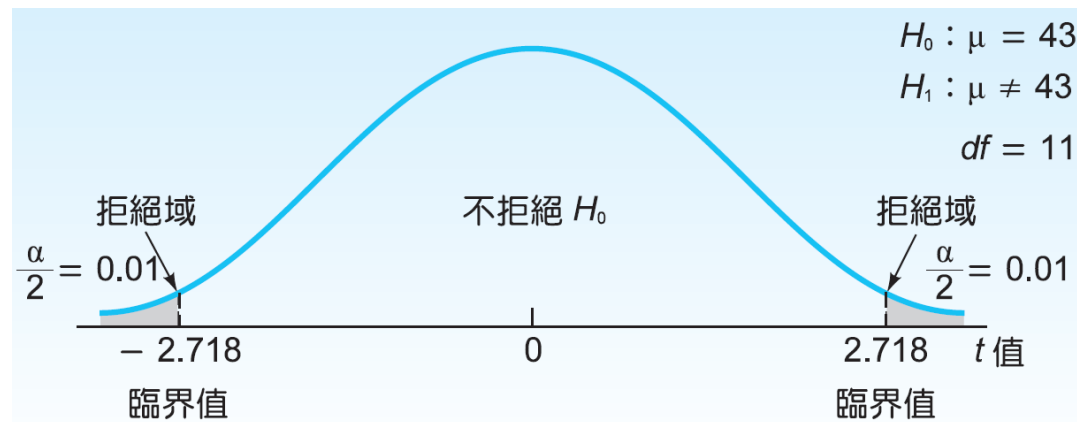
42	39	42	45	43	40	39	41	40	42	43	42
----	----	----	----	----	----	----	----	----	----	----	----

- 可以說平衡錘的平均長度已經改變了嗎？請使用 0.02 顯著水準。

假設檢定－範例六

□ 步驟

- 建立虛無假設與對立假設。 $H_0: \mu = 43$, $H_1: \mu \neq 43$
- 本例為雙尾檢定，自由度是 $n-1 = 12-1 = 11$ 。使用 0.02 顯著水準找出 t 值是 2.718。
- 決策法則：如果檢定統計量 t 落在 -2.718 的左邊或是 2.718 的右邊，則拒絕虛無假設。
- t 分配雙尾檢定的拒絕域，其中 $\alpha = 0.02$



假設檢定－範例六

□ 計算 t 值：

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{41.5 - 43}{1.784/\sqrt{12}} = -2.913$$

□ 落在 -2.718 的左邊區域，拒絕母體平均數是 43 公釐的虛無假設，接受對立假設。

○ 母體平均數不是 43 公釐，機器需再做調整。

□ 比例的定義是成功次數與觀測資料總數量的比例。

○ 計算樣本比例 p 的公式

$$p = \frac{\text{樣本中成功次數}}{\text{樣本數}}$$

$$z = \frac{p - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}}$$

假設檢定－範例六

TABLE 10-2 Calculations of the Sample Standard Deviation

X (mm)	$X - \bar{X}$	$(X - \bar{X})^2$	
42	0.5	0.25	$\bar{X} = \frac{498}{12} = 41.5 \text{ mm}$
39	-2.5	6.25	
42	0.5	0.25	$s = \sqrt{\frac{\sum(X - \bar{X})^2}{n - 1}} = \sqrt{\frac{35.00}{12 - 1}} = 1.784$
45	3.5	12.25	
43	1.5	2.25	
40	-1.5	2.25	
39	-2.5	6.25	
41	-0.5	0.25	
40	-1.5	2.25	
42	0.5	0.25	
43	1.5	2.25	
42	0.5	0.25	
<u>498</u>	<u>0</u>	<u>35.00</u>	

假設檢定－範例七

- ❑ 前次州長選舉顯示，若候選人要當選，至少要在北區贏得80%選票。現任州長想了解繼續連任的機會，計畫在北區隨機抽選2,000位合格選民進行調查。使用假設檢定，了解現任州長連任機會。
- ❑ 州長選舉情形符合二項分配條件，但此範例可使用常態分配去近似二項分配，因 $n\pi$ 與 $n(1-\pi)$ 皆超過5。
 - $n = 2,000$ ， $\pi = 0.8$ （ π 是在本州北區得票率，80%）
 - $n\pi = 2,000 \times 0.8 = 1,600$ 與 $n(1-\pi) = 2,000 \times (1-0.8) = 400$ 都大於5，可以進行母體比例的檢定。

假設檢定－範例七

- ❑ 步驟 1：建立虛無與對立假設。 $H_0: \pi \geq 0.8$ ； $H_1: \pi < 0.8$
- ❑ 步驟 2：選擇顯著水準。為 0.05
- ❑ 步驟 3：選擇檢定統計量。使用 z 統計量，
- ❑ 步驟 4：建立決策法則。
 - 顯著水準 0.05，機率是左尾拒絕域的面積。
 - z 值介於 0 與臨界值間的機率是 $0.5 - 0.05 = 0.45$ 。
 - 查表面積接近 0.45 相對應 z 值 1.65。左尾臨界值是 -1.65。
 - 決策法則：若透過樣本計算出檢定統計量小於 -1.65 時，須拒絕虛無假設，並接受對立假設；反之，則不得拒絕虛無假設 H_0 。

假設檢定－範例七

□ 步驟 5：做決策。

- 抽選 2,000 位合格選民調查，其中 1,550 位選民打算投給現任州長。樣本比例是 $1,550/2,000 = 0.775$ 。

$$z = \frac{\hat{p} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} = \frac{\frac{1,550}{2,000} - 0.8}{\sqrt{\frac{0.8(1-0.8)}{2,000}}} = \frac{0.775 - 0.8}{\sqrt{0.00008}} = -2.8$$

- z 值 -2.8 落在拒絕域中，在顯著水準 0.05 下，拒絕虛無假設。
- 表示樣本比例 77.5% 與假設母體比例 80% 間，2.5% 的差距在統計上是顯著的，此差距可能不是抽樣誤差引起。樣本證據無法支持現任州長會繼續連任。

假設檢定－範例七

- p-值為計算 z 值小於-2.8的機率。
 - 找出介於 0 至 - 2.8 範圍內 z 值的機率是 0.4974。
 - p- 值是 $0.5 - .4974 = 0.0026$ 。此 p- 值小於顯著水準 0.05，表示州長要連任的機率非常小。
 - 顯著水準 0.05 與單尾檢定的拒絕域

