Ch10 假 設 檢 定

Tests of Hypothesis

假設檢定

- 假設檢定是先對母體參數提出假設（主張），然後利用樣本的資訊，再決定是否接受或否決該假設。
  - 某廠牌汽車宣稱『其新推出汽車叫其他廠牌汽車省油 15%。』
  - 某候選人聲稱其支持率高達 50%。
  - 高露潔牙膏公司宣稱每條牙膏重達 200 公克。
  - 某廠商聲稱其產品不良率僅 0.01%。

假設檢定的步驟

- Step 1: 設定虛無假設 (null hypothesis) 與對立假設 (alternative hypothesis)
- Step 2: 選擇檢定統計量(test statistic)
- Step 3: 選擇顯著水準 (level of significance) α 並決定決策法則
- Step 4: 比較樣本統計量與臨界值
- Step 5: 下結論

Step 1:

設定虛無假設 (null hypothesis)

與

對立假設 (alternative hypothesis)

Null Hypothesis (虛無假設)

- Null hypothesis, denoted by H_0, is a statement about the value of a population parameter developed for the purpose of testing numerical evidence.
  - 虛無假設對母體參數提出一個主張，假設此主張為真實（除非能證明此主張非真！）。

5

Alternative Hypothesis (對立假設)

- Alternative hypothesis, denoted by H_1, is a statement that is accepted if the sample data provide sufficient evidence that the null hypothesis is false.
  - 對立假設是相對於虛無假設所提出的另一個不同（相反）的假設或主張，必須有足夠的證據，才能說明此主張為真。

6

Important Things about H_0 and H_1 (1)

- H_0: null hypothesis and H_1: alternate hypothesis
- H_0 and H_1 are mutually exclusive and collectively exhaustive
- H_0 is always presumed to be true
- H_1 has the burden of proof

7

Important Things about H_0 and H_1 (2)

- A random sample is used to "*reject H_0*"
- If we conclude 'do not reject H_0', this does not necessarily mean that the null hypothesis is true, it only suggests that there is not sufficient evidence to reject H_0. If we reject the null hypothesis, then it suggests that the alternative hypothesis may be true.
- Equality is always part of H_0 (e.g. "$=$" , "$\geq$" , "$\leq$"). "$\neq$" "$<$" and "$>$" is always part of H_1

8

How to Set Up a Claim as Hypothesis

- H_0： 現狀、
- H_1： 改變現狀的、犯不得錯的、成本高的

例：

- 法庭上： H_0： 無罪； H_1： 有罪
  （要有足夠證據才能判有罪，否則可能造成一個家庭破碎！）
- 花生工廠： H_0： 一包重**100g**； H_1： 不等於 **100g**
  （要有足夠證據才能說機器有問題，否則需花錢修理！）
- 藥廠發明新藥： H_0： 對人體有害； H_1：對人體無害
  （要有足夠證據才能說新藥不傷身，否則造成身體傷害！）

假設檢定有三種

- 雙尾檢定： H_0： $\theta = \theta_0$； H_1： $\theta \neq \theta_0$
- 左尾檢定： H_0： $\theta \geqq \theta_0$ ； H_1： $\theta < \theta_0$
- 右尾檢定： H_0： $\theta \leqq \theta_0$ ； H_1： $\theta > \theta_0$

Ex. 10.1 汽車耗油的檢定 (1)

台塑汽車生產的迷你小車台塑二號 Matiz 宣稱每公升汽油至少可跑 18.2 公里。這對於許多自小客車駕駛人來說，很有吸引力。假設張先生想買一部 Matiz，但又恐不是真的那麼省油。因而向購買 Matiz 的朋友打聽，以檢驗 Matiz 每公升汽油至少可跑 18.2 公里是否正確。請幫他設立假設檢定的兩個假設。

Ex. 10.1 汽車耗油的檢定 (1)

假設製造商的宣稱是真實的，故虛無假設為每公升可行駛至少 18.2 公里。樣本資料必須有足夠證據，方能證明製造商的宣稱『每公升汽油可行駛18.2公里以上』為假，因此，對立假設設為每公升汽油行駛里程小於 18.2 公里，亦即:

$$H_0 : \mu \geq 18.2$$

$$H_1 : \mu < 18.2$$

Ex. 10.2 吸菸比例是否提高？(1)

根據衛生署國民健康局的調查顯示，年青人吸菸的比例有提高的趨勢。最近調查顯示，男性18~21歲開始吸菸的比例已超過 20%，某研究員想檢定這項調查結果是否屬實，請幫他設立假設檢定。

Ex. 10.2 吸菸比例是否提高？(2)

18~21 歲開始吸菸比例小於等於 20% 設為虛無假設，而吸菸比例大於 20% 設為對例假設，即：

$$H_0 : p \le 20\%$$

$$H_1 : p > 20\%$$

Ex. 10.3 機器運作正常嗎？

豐正馬達工廠在正常運作時產品的瑕疵率為 0.02，但若運作不正常，則瑕疵率將達 0.04。現經理懷疑工廠運作不正常，擬進行假設檢定，問如何設立假設？

$$H_0 : p = 0.02$$
$$H_1 : p = 0.04$$

假設檢定的兩種錯誤 (1)

- 型 I 誤差 (Type I Error)：
  --- **當 H_0 為真，而拒絕 H_0 所發生的錯誤。**
- 型 II 誤差 (Type II Error)：
  --- **當 H_0 為假，而不拒絕 H_0 所發生的錯誤（或 H_1 為真，沒有接受 H_1 為真所發生的錯誤）。**

假設檢定的兩種錯誤 (2)

Null Hypothesis	Researcher: Does Not Reject H_0	Researcher: Rejects H_0
H_0 is true	Correct decision	Type I error
H_0 is false	Type II error	Correct decision

17

例：法官判案

- H_0： 無罪 ； H_1：有罪
-

		真實情形	
		H_0： 無罪	H_1：有罪
決策	不拒絕 H_0（判無罪）	正確	型 II 錯誤
	拒絕 H_0（判有罪）	型 I 錯誤	正確

18

兩種錯誤的發生機率 (1)

- 型 I 錯誤發生的機率 α:

$$\alpha = \mathbf{P(拒絕\ H_0 \mid H_0\ 為真)}$$

又稱為 α 錯誤或 α 風險，或顯著水準 (Significance Level)。

- 型 II 錯誤發生的機率 β:

$$\beta = \mathbf{P(不拒絕\ H_0 \mid H_1\ 為真)}$$

又稱為 β 錯誤或 β 風險。

19

兩種錯誤的發生機率 (2)

- α 與 β 二機率互為消長。
  - $\alpha\uparrow \Rightarrow \beta\downarrow$ ； $\alpha\downarrow \Rightarrow \beta\uparrow$
  - $\alpha + \beta \neq 1$
- H_0: 無罪 ； H_1: 有罪
  - α = P(無罪判成有罪)
  - ➔ 造成Type I error的原因在於判有罪，所以都判無罪，如此可以使 $\alpha\downarrow$，然而卻亦造成有罪的也被判成無罪，因此 $\beta\uparrow$
  - β = P(有罪判成無罪)
  - ➔ 造成Type II error的原因在於判無罪，所以都判有罪，如此可以使 $\beta\downarrow$，然而卻亦造成無罪的也被判成有罪，因此 $\alpha\uparrow$

20

兩種錯誤的發生機率 (2)

- 假設檢定發生型 I 或型 II 錯誤，均可能導致損失，因此希望α與β不要太大。但在相同樣本數下，無法同時使α與β變小。
- 一般認為型 I 錯誤的後果較嚴重，因此希望α較小，換言之，希望證據相當充足的情形下，才推翻 H_0。
- α越小，表示對H_1的檢驗越嚴格，不輕易接受 H_1 。
- 常用的 α 值為 0.01, 0.05 與 0.1 。

21

Ex.10.4 生態環境的保護

經濟發展的代價之一是環境汙染。像工廠排放的廢水、廢氣，汽車機車排放的廢氣，家庭棄置之廢棄物等所造成的污染。當所得水準相對低的時候，人們甚少注意到環境汙染的問題，然而隨著所得的提高，人們逐漸發現環境保護的重要性不亞於所得的提高，因此，對各種汙染的管制檢查日趨嚴格。例如對廢水中含油量θ的檢定，排放廢氣量θ的檢定，可假設為：

$H_0: \theta \geq \theta_0$ (廢氣排放量超過規定)

$H_1: \theta < \theta_0$ (廢氣排放量合乎規定)

此時檢驗人員將α值設得相當小，換句話說，檢驗人員不希望輕易支持不合格的廠商的汙染行為，以達到生態環境保護的目的。

22

Step 2:

選擇檢定統計量(test statistic)

23

檢定統計量 (Test Statistic)

- 檢定統計量是由樣本所算出來的一個值，用來決定是否接受或拒絕 H_0。
- 不同參數的假設檢定，使用不同的檢定統量。
- 常用的檢定統計量有：Z, t, F 與 χ^2。

24

檢定統計量 (Test Statistic)

- 在母體平均數的假設檢定裡，不同的情形下使用不同的檢定統計量。

1. 母體 σ^2 已知，無論樣本數大小，皆使用常態分配: $Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}}$

2. 母體 σ^2 未知：

   (a) 當樣本數 n ≥ 30，可以使用常態分配，但 σ 用 s 取代: $Z = \dfrac{\overline{X} - \mu}{s / \sqrt{n}}$

   (b) 當樣本數 n < 30，使用 t 分配: $t = \dfrac{\overline{X} - \mu}{s / \sqrt{n}}$ (自由度 $n-1$)

25

Step 3:

選擇顯著水準 (level of significance) α
並決定決策法則

26

決定決策法則 (1)

- 研究人員必須決定一個決策法則，以瞭解何時『**不拒絕**』 H_0 ；何時拒絕 H_0 。
  - 一般我們說『**不拒絕**』H_0，而不說接受 H_0，因為我們只是沒有足夠證據拒絕，而不是接受。
- 決策法則通常是決定一個不拒絕域 (Nonrejection Region, 或稱接受域) 與拒絕域 (Rejection Region)。當檢定統計量落入：
  - 不拒絕域：『**不拒絕**』 H_0 ！
  - 拒絕域：拒絕 H_0 ；接受 H_1 ！

27

決定決策法則 (2)

- 接受域與拒絕域的接點，稱為臨界點(Critical Point)。
- 臨界值的決定，是根據顯著水準 α 並利用機率分配計算而得，分成三種形式：
  - 雙尾檢定 (Two-tail Test)
  - 右尾檢定 (Right-tail Test)
  - 左尾檢定 (Left-tail Test)

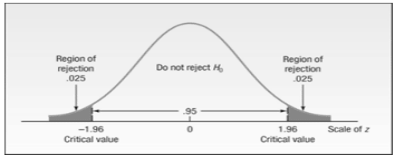
28

雙尾檢定 (Two-tail Test) (1)

$H_0 : \mu = \text{value}$

$H_1 : \mu \neq \text{value}$

Reject H_0 if



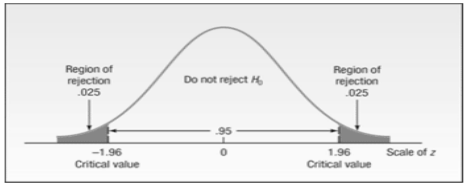
$$|z| = \left| \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \right| > Z_{\alpha/2} \quad (\ \sigma \text{ is known !})$$

$$|t| = \left| \frac{\overline{X} - \mu}{s / \sqrt{n}} \right| > t_{\alpha/2,\, n-1} \quad (\ \sigma \text{ is unknown !})$$

29

雙尾檢定 (Two-tail Test) (2)

- 臨界點：$\pm Z_{\alpha/2}$ (σ is known); $\pm t_{\alpha/2}$ (σ is unknown).
- 接受域：$[-Z_{\alpha/2}, Z_{\alpha/2}]$; $[-t_{\alpha/2}, t_{\alpha/2}]$
- 拒絕域： |檢定統計量| > $|Z_{\alpha/2}|$ 或 $|t_{\alpha/2}|$



30

雙尾檢定 (Two-tail Test) (3)

- α=0.05，接受域$[-Z_{\alpha/2}, Z_{\alpha/2}]$= [-1.96, 1.96]
- α=0.1，接受域 $[-Z_{\alpha/2}, Z_{\alpha/2}]$=[-1.645, 1.645]
- α=0.05， n=20 且σ未知，則接受域為 $[-t_{\alpha/2,n-1}, t_{\alpha/2,n-1}]$ = [-2.093, 2.093]
- α=0.01， n=15 且σ未知，則接受域為 $[-t_{\alpha/2,n-1}, t_{\alpha/2,n-1}]$ = [-2.977, 2.977]

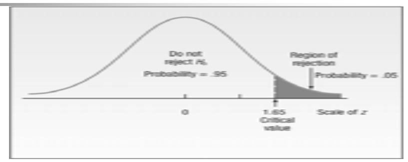
31

右尾檢定 (Right-tail Test) (1)

$H_0 : \mu \leq \text{value}$

$H_1 : \mu > \text{value}$

Reject H_0 if



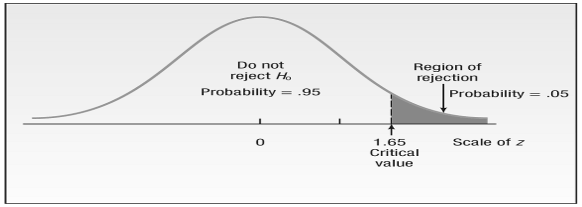
$$z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} > Z_{\alpha} \quad (\ \sigma \text{ is known !})$$

$$t = \frac{\overline{X} - \mu}{s / \sqrt{n}} > t_{\alpha,\, n-1} \quad (\ \sigma \text{ is unknown !})$$

32

右尾檢定 (Right-tail Test) (2)

- 臨界點： Z_α (σ is known); t_α (σ is unknown).
- 接受域：[-∞ , Z_α]; [-∞, t_α]
- 拒絕域： 檢定統計量 > Z_α 或 t_α



33

右尾檢定 (Right-tail Test)(3)

- α=0.05，接受域 [-∞ , Z_α] = [-∞ , 1.96]
- α=0.1，接受域 [-∞ , Z_α] =[-∞ , 1.645]
- α=0.05， n=20 且 σ 未知，則接受域為 [-∞ , $t_{\alpha,n-1}$] = [-∞ , 2.093]
- α=0.01， n=15 且 σ 未知，則接受域為 [-∞, $t_{\alpha/,n-1}$] = [-∞ , 2.977]

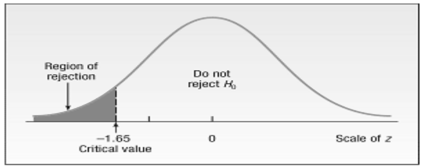
34

左尾檢定 (Left-tail Test) (1)

$H_0: \mu \geq$ value

$H_1: \mu <$ value

Reject H_0 if



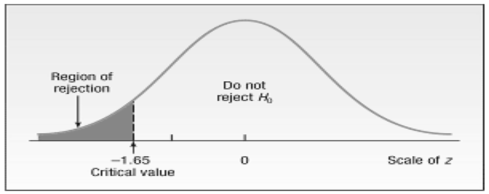
$$z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} < -Z_\alpha \quad (\sigma \text{ is known !})$$

$$t = \frac{\overline{X} - \mu}{s / \sqrt{n}} < -t_{\alpha,\, n-1} \quad (\sigma \text{ is unknown !})$$

35

左尾檢定 (Left-tail Test) (2)

- 臨界點： $Z_{-\alpha}$ (σ is known); $t_{-\alpha}$ (σ is unknown).
- 接受域：[$Z_{-\alpha}$, ∞]; [$t_{-\alpha}$, -∞]
- 拒絕域： 檢定統計量 < $Z_{-\alpha}$ 或 $t_{-\alpha}$



36

左尾檢定 (Right-tail Test)(3)

- α=0.05，接受域 [$Z_{-\alpha}$, ∞] = [-1.96, ∞]
- α=0.1，接受域 [$Z_{-\alpha}$, ∞] = [-1.645, ∞]
- α=0.05， n=20 且 σ 未知，則接受域為
  [$t_{-\alpha,n-1}$, ∞] = [-2.093, ∞]
- α=0.01， n=15 且 σ 未知，則接受域為
  [$t_{-\alpha,n-1}$, ∞] = [-2.977, ∞]

37

Step 4:

結　論

38

結論

- 檢定的結果需說明檢驗的結果，而非只描述接受或拒絕 H_0。
- 結果若是接受 H_0，不表示 H_0 為真，僅是樣本沒有足夠證據推翻他。
- 結果若是拒絕 H_0，即表示樣本提供足夠證據接受 H_1 為真。

39

Ex. 10.5 營業額的檢定 – 該不該調房租 (1)

大統超商與其房東賴先生因為房屋租金產生歧見，而無法簽定續租合約。大統公司的翁經理辯稱，因為顧客不多，生意不好，故不同意房東依物價指數調整房租的要求；房東則認為該公司生意不錯，翁經理的理由不足採信。雙方各執一詞。最後，他們同意請陳教授來仲裁。陳教授建議就營業額的多寡作為租金調整與否的依據，雙方同意若每日營業收入在 100,000 以上，則應調整房租。問陳教授怎樣才能做出公平的仲裁呢？

40

Ex. 10.5 營業額的檢定 – 該不該調房租 (2)

首先陳教授隨機抽取大統過去 64 天的營業發票存根，計算得出每日營業額的平均數 $\overline{X}=98$ (千元)，標準差 $S=10$ (千元)。接著對營業額進行假設檢定，檢定的步驟為：

(1) 設立如下兩個假設：

$H_0: \mu \geq 100$ 千元 (營業收入大於等於十萬元)

$H_1: \mu < 100$ 千元 (營業收入小於十萬元)

(2) 選擇假設統計量

以 $\overline{X}$ 作為檢定統計量。因為 $n=64$ 為大樣本，而 H_0 為真的情況下，$\overline{X}$ 的抽樣分配近似常態分配。

$$\overline{X} \sim N(100,(10)^2/64)$$

Ex. 10.5 營業額的檢定 – 該不該調房租 (3)

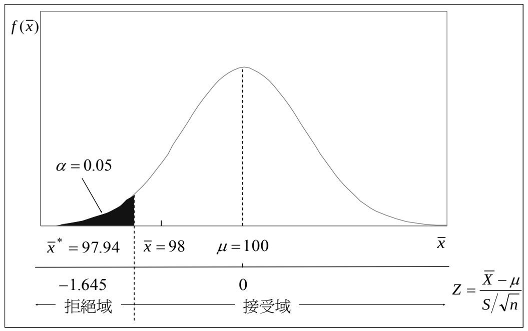
(3) 選擇顯著水準及決定決策法則

居於公正客觀的立場，陳教授選擇顯著水準 (控制型 I 錯誤的機率) $\alpha=5\%$，因對立假設為＜，故採左尾檢定。接著他計算臨界值 (決定拒絕域與接受域)： $\overline{X}^*=\mu_0-Z_\alpha S_{\overline{X}}=100-Z_{0.05}\dfrac{S}{\sqrt{n}}$

$$=100-Z_{0.05}\frac{10}{\sqrt{64}}=100-1.645\times\frac{10}{8}=97.94$$

決策規則為：若 $\overline{X}<97.94$ 則拒絕 H_0， $\overline{X}\geq 97.94$ 則接受 H_0。

Ex. 10.5 營業額的檢定 – 該不該調房租 (4)



Ex. 10.5 營業額的檢定 – 該不該調房租 (5)

(4) 比較樣本平均數的觀察值與臨界值

樣本平均數的觀察值 $\overline{X}=98$ 大於臨界值 97.94，落於接受域內。

(5) 陳教授於是根據檢定的結果，提出檢定報告如下：「在顯著水準 $\alpha=5\%$ 下，根據 64 天的營業資料作營業額小於 100,000 的假設檢定，檢定結果接受虛無假設。」根據雙方約定，大統公司應同意調整房租。

應注意的是，上面的結論是在顯著水準 $\alpha=5\%$ 下的結果。當 α 顯著水準不同時，可能會有不同的結論，因此下結論時必須說出顯著水準 α 值。

Ex. 10.5 營業額的檢定 – 該不該調房租 (5)

在上例中，如果陳教授將顯著水準定為 10% 時，結論是否和顯著水準 5% 時相同呢？

45

Ex. 10.5 營業額的檢定 – 該不該調房租 (6)

在顯著水準 $\alpha = 10\%$ 時，臨界值為：

$$\overline{X}^* = 100 - 1.28 \times \frac{10}{\sqrt{64}} = 98.4$$

由於樣本平均數得觀察值 $\overline{X} = 98$ 小於臨界值 98.4，落在拒絕域，因此拒絕虛無假設。

由此可知，顯著水準不同，可能有不同結論。

46

Ex. 10.7 機器性能的檢定 (1)

- 日裕精密儀器公司原有的工作母機平均每小時生產零件 490 件，但已經使用十多年了。現該公司準備更新機器。現有大新機器公司宣稱其所製造的工作母機每小時至少可生產零件 600 件，但價格高出 5%。採購部門要求廠商部門檢定大新的機器是否真的如其所宣稱，以做為採購的參考。問廠務部門的王副理如何檢定呢？

47

Ex. 10.7 機器性能的檢定 (2)

王副理首先隨機取得36部大新的機器，並計算36部機器每小時生產零件的平均數，結果得出$\overline{X}$ =605，標準差S=12，。接著進行檢定：

1.設立兩個假設：H_0: $\mu \le 600$ & H_1: $\mu > 600$

2.選擇檢定統計量

以$\overline{X}$做爲檢定統計量。因爲n=36爲大樣本，在H_0爲真的情況下，

$\overline{X}$的抽樣分配近似常態分配 $\overline{X} \sim N(600,12^2/36)$

48

Ex. 10.7 機器性能的檢定 (3)

3.選擇顯著水準α=5%，因對立假設爲>的符號，故採右尾檢定計算

臨界值 (決定拒絕域與接受域)：

$$\bar{X}^* = 600 + Z_{0.05}\frac{12}{\sqrt{36}} = 600 + 1.645 \times 2 = 603.29$$

決策法則爲：若$\bar{X} > 603.29$則拒絕H_0。若$\bar{X} \le 603.29$則接受H_0。

Ex. 10.7 機器性能的檢定 (4)

4.比較檢定統計量及臨界值。

由於檢定統計量的觀察值$\bar{X}$大於臨界值603.29，落在拒絕域。

5.下結論

王副理根據檢定結果可以下結論：在α=0.05下，拒絕H_0。

亦即，在α=0.05下，大新的機器性能如其所稱，每小時生產600個零件

Ex. 10.8 (1)

- 最新式的筆記型電腦的液晶顯示器標準長度為 10 英吋。太長或太短都不合用。現康定電腦公司購進 49000 個顯示板，品管部門周小姐奉命做檢定。問周小姐應如何進行？

Ex. 10.8 (2)

- 這是一個雙尾的檢定問題。因為液晶板的標準為 10 英吋，不能長於或短於 10 英吋。首先周小姐從貨品中取出49 個液晶螢幕為一組樣本，並計算得出樣本平均數為 10.09，標準差為 0.28，接著進行檢定工作。

Ex. 10.8 (3)

1.設立兩個假設

H_0: μ=10　H_1: $\mu \neq 10$

2.選擇檢定統計量

以$\bar{X}$做爲檢定統計量。因爲n=49爲大樣本， 在 H_0 爲真的情況下，$\bar{X}$的抽樣分配近似常態分配 $\bar{X} \sim N(10, 0.28^2/49)$

3.選擇顯著水準即決定決策法則

選擇α=0.1，因對立假設有≠的符號，故採雙尾檢定，拒絕域在 $\bar{X}$ 抽樣分配的兩尾。因爲 α=0.1，查標準常態機率表知，臨界值 $-Z_{0.05}$=1.645, $Z_{0.05}$=1.645。

53

Ex. 10.8 (4)

接著計算上臨界值與下臨界值

$$\bar{X}_L^* = \mu - Z_{\alpha/2}\frac{S}{\sqrt{n}} = 10 - 1.645 \times \frac{0.28}{\sqrt{49}} = 10 - 0.0658 = 9.9342$$

$$\bar{X}_U^* = \mu + Z_{\alpha/2}\frac{S}{\sqrt{n}} = 10 + 1.645 \times \frac{0.28}{\sqrt{49}} = 10 + 0.0658 = 10.0658$$

4. 比較檢定統計量的樣本觀察值與臨界值

樣本檢定統計量的觀察值$\bar{X}$=10.09大於臨界值的上限10.066，落在拒絕域。

5. 下結論

因樣本檢定統計量的觀察值落在拒絕域。故周小姐可下結論：

在α=0.1水準下，液晶顯示板長度不爲10 英吋，貨品不能接受。

54

Ex. 10. 9 Testing for a Population Mean with a Known Population Standard Deviation (1)

Jamestown Steel Company manufactures and assembles desks and other office equipment at several plants in western New York State. The weekly production of the Model A325 desk at the Fredonia Plant follows the normal probability distribution with a mean of 200 and a standard deviation of 16. Recently, because of market expansion, new production methods have been introduced and new employees hired. The vice president of manufacturing would like to investigate whether there has been a *change* in the weekly production of the Model A325 desk.

55

Ex. 10. 9 Testing for a Population Mean with a Known Population Standard Deviation (2)

Step 1: State the null hypothesis and the alternate hypothesis.

H_0: μ = 200

H_1: μ ≠ 200

(note: keyword in the problem "has changed")

Step 2: Select the level of significance.

α = 0.01 as stated in the problem

Step 3: Select the test statistic.

Use Z-distribution since σ is known

56

Ex. 10. 9 Testing for a Population Mean with a Known Population Standard Deviation (3)

Step 4: Formulate the decision rule.

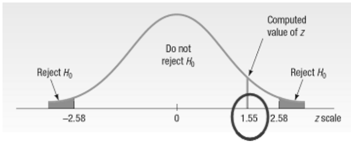
Reject H_0 if $|Z| > Z_{\alpha/2}$

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

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

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

1.55 is not > 2.58



Step 5: Make a decision and interpret the result.

Because 1.55 does not fall in the rejection region, H_0 is not rejected. We conclude that the population mean is not different from 200. So we would report to the vice president of manufacturing that the sample evidence does not show that the production rate at the Fredonia Plant has changed from 200 per week.

Ex. 10. 10 Testing for a Population Mean with a Known Population Standard Deviation (1)

Suppose in the previous problem the vice president wants to know whether there has been an *increase* in the number of units assembled. To put it another way, can we conclude, because of the improved production methods, that the mean number of desks assembled in the last 50 weeks was more than 200?

Recall: σ =16, n=50, α =.01

Ex. 10. 10 Testing for a Population Mean with a Known Population Standard Deviation (2)

Step 1: State the null hypothesis and the alternate hypothesis.

H_0: $\mu \leq 200$

H_1: $\mu > 200$

(note: keyword in the problem "an *increase*")

Step 2: Select the level of significance.

α **= 0.01 as stated in the problem**

Step 3: Select the test statistic.

Use Z-distribution since σ is known

Ex. 10. 10 Testing for a Population Mean with a Known Population Standard Deviation (3)

Step 4: Formulate the decision rule.

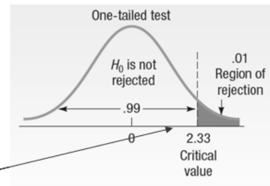
Reject H_0 if $Z > Z_\alpha$

$$Z > Z_\alpha$$

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

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

1.55 is not > 2.33



Ex. 10. 10 Testing for a Population Mean with a Known Population Standard Deviation (4)

Step 5: Make a decision and interpret the result.

Because 1.55 does not fall in the rejection region, H_0 is not rejected. We conclude that the average number of desks assembled in the last 50 weeks is not more than 200

61

Ex. 10. 11 Testing for the Population Mean: Population Standard Deviation Unknown (1)

The McFarland Insurance Company Claims Department reports the mean cost to process a claim is $60. An industry comparison showed this amount to be larger than most other insurance companies, so the company instituted cost-cutting measures. To evaluate the effect of the cost-cutting measures, the Supervisor of the Claims Department selected a random sample of 26 claims processed last month. The sample information is reported below.

At the 0.01 significance level is it reasonable a claim is *now less than $60*?

$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				

62

Ex. 10. 11 Testing for the Population Mean: Population Standard Deviation Unknown (2)

Step 1: State the null hypothesis and the alternate hypothesis.

H_0: $\mu \geq \$60$

H_1: $\mu < \$60$

(note: keyword in the problem "*now less than*")

Step 2: Select the level of significance.

α = 0.01 as stated in the problem

Step 3: Select the test statistic.

Use *t*-distribution since σ is unknown

63

t-Distribution Table (portion)

TABLE 10–1 A Portion of the *t* Distribution Table

	Confidence Intervals					
	80%	90%	95%	98%	99%	99.9%
	Level of Significance for One-Tailed Test, α					
df	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	$\boxed{2.485}$	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

64

Ex. 10. 11 Testing for the Population Mean: Population Standard Deviation Unknown (3)

Step 4: Formulate the decision rule.

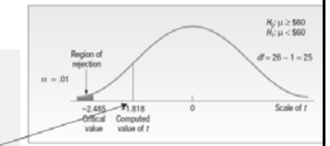
Reject H_0 if $t < -t_{\alpha,n-1}$

$$t < -t_{\alpha,n-1}$$

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} < -t_{\alpha,n-1}$$

$$\frac{\$56.42-\$60}{\$10.04/\sqrt{26}} < -t_{.01,26-1}$$

$$-1.818 \text{ is not } < -2.485$$



Step 5: Make a decision and interpret the result.

Because -1.818 does not fall in the rejection region, H_0 is not rejected at the .01 significance level. We have not demonstrated that the cost-cutting measures reduced the mean cost per claim to less than $60. The difference of $3.58 ($56.42 - $60) between the sample mean and the population mean could be due to sampling error.

Ex. 10. 12 Testing for a Population Mean with an Unknown Population Standard Deviation (1)

The current rate for producing 5 amp fuses at Neary Electric Co. is 250 per hour. A new machine has been purchased and installed that, according to the supplier, will increase the production rate. A sample of 10 randomly selected hours from last month revealed the mean hourly production on the new machine was 256 units, with a sample standard deviation of 6 per hour.

At the 0.05 significance level can Neary conclude that the new machine is faster?

Ex. 10. 12 Testing for a Population Mean with an Unknown Population Standard Deviation (2)

Step 1: State the null and the alternate hypothesis.

H_0: $\mu \leq 250$; H_1: $\mu > 250$

Step 2: Select the level of significance.

It is .05.

Step 3: Find a test statistic. Use the t distribution because the population standard deviation is not known and the sample size is less than 30.

Ex. 10. 12 Testing for a Population Mean with an Unknown Population Standard Deviation (3)

Step 4: State the decision rule.

There are 10 – 1 = 9 degrees of freedom.

The null hypothesis is rejected if $t > 1.833$.

$$t = \frac{\bar{X}-\mu}{s/\sqrt{n}} = \frac{256-250}{6/\sqrt{10}} = 3.162$$

Step 5: Make a decision and interpret the results.

The null hypothesis is rejected. The mean number produced is more than 250 per hour.

Ex. 10.13 (1)

- 台銀經理感覺客戶有減少的跡象。調查發現客戶抱怨等待時間太長，平均每人等待15分鐘。為改善此一現象，重新調整櫃檯人員。現隨機抽取 16 個客戶，得平均等待11分鐘，標準差2分鐘。問在α=0.05下，平均等待時間在何範圍內才能顯示效率有改善？或該措施有效？

69

Ex. 10.13 (2)

1.設立兩假設

想檢定此措施是否有效提高效率，若等待時間縮短表示有效，故兩假設爲: H_0: $\mu \geq 15$; $H_1<15$

2.選擇檢定統計量

小樣本 n=16，母體爲常態且母體變異數未知，故選擇 t 分配來檢定母體平均數。

70

Ex. 10.13 (3)

3. 根據顯著水準並決定拒絕域及接受域

α=0.05 下，因對立假設爲 <，所以爲左尾檢定。

自由度 df = 16-1 = 15，故臨界值 $-t_{15,005} = -1.753$

$\bar{X}^* = 15\text{-}t_{15,005} S/\sqrt{n} = 15 - 1.753 * 2/\sqrt{16} = 14.124$

亦即若 $\bar{X} \leq 14.124$ 則拒絕 H_0 ，而認爲效率有提高。

4. 下結論

因 $\bar{X}$=11 分鐘落入拒絕域，因此拒絕H_0。結論爲在 α=0.05 下，台銀調整服務人員的措施有效。

71

母體比例的假設檢定

Test of Hypothesis for Population Mean

72

母體比例的假設檢定

- 母體比例經常是我們想要瞭解的母體參數之疑。例如：
  - 某參選人的支持率。
  - 產品的良率。
  - 產品的市佔率。
- 除了可以使用統計估計去探討母體比例，亦可以使用假設檢定的方式去推論母體比例。

Assumptions in Testing a Population Proportion using the z-Distribution (1)

- A random sample is chosen from the population.
- It is assumed that the binomial assumptions are met:
  - the sample data collected are the result of counts
  - the outcome of an experiment is classified into one of two mutually exclusive categories—a "success" or a "failure"
  - the probability of a success is the same for each trial
  - the trials are independent

Assumptions in Testing a Population Proportion using the z-Distribution (2)

- The test we will conduct shortly is appropriate when both $n\pi$ and $n(1-\pi)$ are at least 5.
- When the above conditions are met, the normal distribution can be used as an approximation to the binomial distribution

Test Statistic for a Single Population Proportion

$$X \sim \text{Bernoulli}(\pi) \Rightarrow \sum X \sim Binomial(n,\pi)$$

$$\text{CLT} \Rightarrow \text{P} = \frac{\sum X}{n} = \overline{X} \approx \text{N}\left(\text{E[X]}, \frac{\sigma_X^2}{\text{n}^2}\right) = N\left(\pi, \frac{\pi(1-\pi)}{n}\right)$$

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

Hypothesis Setups for Testing a Proportion (π)

H_0: π = value	H_0: $\pi \geq$ value	H_0: $\pi \leq$ value
H_1: $\pi \neq$ value	H_1: $\pi <$ value	H_1: $\pi >$ value
Reject H_0 if: $\lvert Z \rvert > Z_{\alpha/2}$	Reject H_0 if: $Z < -Z_\alpha$	Reject H_0 if: $Z > Z_\alpha$

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

77

Ex. 10.14 Test Statistic for Testing a Single Population Proportion (1)

Suppose prior elections in a certain state indicated it is necessary for a candidate for governor to receive at least 80 percent of the vote in the northern section of the state to be elected. The incumbent governor is interested in assessing his chances of returning to office and plans to conduct a survey of 2,000 registered voters in the northern section of the state. Using the hypothesis-testing procedure, assess the governor's chances of reelection.

78

Ex. 10.14 Test Statistic for Testing a Single Population Proportion (2)

Step 1: State the null hypothesis and the alternate hypothesis.

H_0: $\pi \geq .80$

H_1: $\pi < .80$

(note: keyword in the problem "*at least*")

Step 2: Select the level of significance.

α = 0.01 as stated in the problem

Step 3: Select the test statistic.

Use Z-distribution since the assumptions are met and $n\pi$ and $n(1-\pi) \geq 5$

79

Ex. 10.14 Test Statistic for Testing a Single Population Proportion (3)

Step 4: Formulate the decision rule.

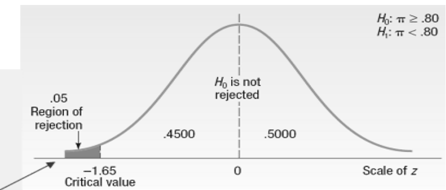
Reject H_0 if $Z < -Z_\alpha$

$$Z < -Z_\alpha$$

$$\frac{p-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} < -z_\alpha$$

$$\frac{\frac{1{,}550}{2{,}000} - .80}{\sqrt{\frac{.80(1-.80)}{2{,}000}}} < -1.65$$

$$-2.80 < -1.65$$



80

Ex. 10.14 Test Statistic for Testing a Single Population Proportion (4)

Step 5: Make a decision and interpret the result.

The computed value of *z* (2.80) is in the rejection region, so the null hypothesis is rejected at the .05 level. The difference of 2.5 percentage points between the sample percent (77.5 percent) and the hypothesized population percent (80) is statistically significant. The evidence at this point does not support the claim that the incumbent governor will return to the governor's mansion for another four years.

81

P-Value

82

P –Value (1)

- 假設檢定必須先設定顯著水準 α，才能算臨界值，進而判斷是否拒絕或接受。
- 不同的值 α，就可能有不同的結果，因此很難決定 α。
  - 例如： $\alpha = 0.05$ 不拒絕 H_0； $\alpha = 0.049$ 拒絕 H_0。
- 因此，統計人員在執行統計檢定時，提供 P-Value，將拒絕或不拒絕 H_0留給決策者自行決定。
- P-Value 是以檢定統計量為臨界值的顯著水準 α。

83

P –Value：雙尾檢定

$H_0 : \mu = \mu_0$

$H_1 : \mu \neq \mu_0$

$$\text{P-value} = \begin{cases} 2 \times P\left(\bar{X} \geq \bar{X}_0 \mid H_0 : \mu = \mu_0\right) & \text{, if } \bar{X}_0 > \mu_0 \\ 2 \times P\left(\bar{X} \leq \bar{X}_0 \mid H_0 : \mu = \mu_0\right) & \text{, if } \bar{X}_0 < \mu_0 \end{cases}$$

84

P –Value：右尾檢定

$H_0 : \mu \le \mu_0$

$H_1 : \mu > \mu_0$

$\text{P-value} = P\left(\bar{X} \ge \bar{X}_0 \mid H_0 : \mu = \mu_0\right)$

P –Value：左尾檢定

$H_0 : \mu \ge \mu_0$

$H_1 : \mu < \mu_0$

$\text{P-value} = P\left(\bar{X} \le \bar{X}_0 \mid H_0 : \mu = \mu_0\right)$

P-Value 檢定的決策法則

- 若 P-Value ≧ α ，則接受虛無假設。
- 若 P-Value＜α ，則拒絕虛無假設。

Ex. 10.15 (1)

- 根據主計處調查 91 年台灣地區每人每年健保支出平均為 2.6 萬。今有一研究人員懷疑此一結果，她隨機抽取 25 人得平均為 2.4 萬元，標準差為 4.8 萬元，問在常態母體分配的假設下，平均醫療支出是否為 2.6 萬元？分別用臨界值法、t值法與P值法來做檢定。 α=0.05。

Ex. 10.15 (2)

由題意知 $n = 25$ $\bar{X}$=2.4 S=4.8

1.設立兩個假設

H_0: μ=2.6 H_1: $\mu \neq 2.6$

2.選擇檢定統計量

小樣本 n=25，母體爲常態且母體變異數未知故以 t 檢定統計量來檢定母體平均數。

3.根據顯著水準決定拒絕域或接受域

α=0.05，對立假設爲 ≠，所以採雙尾檢定，每尾爲 α/2=0.025。拒絕域在 t 分配的兩尾。

89

Ex. 10.15 (3)

自由度 df = n-1 = 24，查 t 值表可知 : $t_{0.025} = 2.064$

計算$\bar{X}$的臨界值

$$\bar{X}_L^* = 2.6 - 2.064\frac{4.8}{\sqrt{25}} = 0.619$$

$$\bar{X}_H^* = 2.6 + 2.064\frac{4.8}{\sqrt{25}} = 4.581$$

若 $0.619 \le \bar{X} \le 4.581$ 則接受 H_0

4.計算檢定統計量或 P 值

t 統計量爲 $t = \dfrac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} = \dfrac{2.4-2.6}{\frac{4.8}{\sqrt{25}}} = -0.208 > t_{24,0.025} = -2.064$

90

Ex. 10.15 (4)

計算 P 值

$$P值 = 2*P(\bar{X} \le 2.4|\mu=2.6) = 2*P(\frac{\bar{X}-\mu_0}{S/\sqrt{n}} \ge \frac{2.4-2.6}{4.8/\sqrt{25}})$$

$$= 2*P(t_{24} \le -0.208) = 2*0.4185 = 0.837 > \alpha = 0.05$$

5.下結論

因 $\bar{X}$=2.4 落再接受域且檢定統計量 t=0.208 落在接受域，故不拒絕虛無假設。結論爲：在 α=0.05 下，每人每年的醫療支出約爲2.6萬。

91

Ex. 10.16 (1)

- 台新白米廠每包白米每包標示重量為 3 公斤。其經銷商廣信進貨1000 包，經抽取 49 包，結果得出樣本平均數為 2.95，標準差 S = 0.38。在 α=0.05 下，若以 P 值法檢定，問這批白米是否可以接受？

92

Ex. 10.16 (2)

由題意知，n=49 $\bar{X}$=2.95 S=0.38 α=0.05

1.設立兩個假設

H_0: $\mu \geq 3$公斤 H_1: $\mu<3$公斤

2.選擇檢定統計量

以 $\bar{X}$ 為檢定統計量，因為 n=49 為大樣本，在 H_0 為真的情況下，$\bar{X}$ 的抽樣分配近似常態分配 $\bar{X} \sim N(3, \frac{0.38^2}{49})$。

93

Ex. 10.16 (3)

3.計算P值

對立假設的符號為<，故應是左尾檢定。此時 P值為 $\bar{X} \leq 2.95$ 公斤的機率，即

$$P值 = P(\bar{X} \leq \bar{X}_0|\mu=3) = P(\bar{X} \leq 2.95|\mu=3) = P(\frac{\bar{X}-3}{0.0542} \leq \frac{2.95-3}{0.0542})$$

$$= P(Z \leq -0.92) = 0.1788$$

4.下結論

由於 P = 0.1788 > α = 0.05，因此不拒絕虛無假設，即台新白米重量符合標示，該批貨品可以接受。

94

Ex. 10.17 (1)

- 根據主計處調查91年台灣地區每人每年健保支出平均為 2.6 萬。今有一研究人員懷疑此一結果，她隨機抽取 25 人得平均為 2.4 萬元，標準差為 4.8 萬元，問在常態母體分配的假設下，平均醫療支出是否為 2.6 萬元？分別用臨界值法、t 值法與 P 值法來做檢定。 α=0.05

95

Ex. 10.17 (2)

由題意知 $n = 25$ $\bar{X}$=2.4 S=4.8

1.設立兩個假設: H_0: μ=2.6; H_1: $\mu \neq 2.6$

2.選擇檢定統計量

小樣本 n=25，母體為常態且母體變異數未知故以 t 檢定統計量來檢定母體平均數。

3.根據顯著水準決定拒絕域或接受域

α=0.05，對立假設為 ≠，所以採雙尾檢定，每尾為 α/2=0.025。拒絕域在 t 分配的兩尾。

96

Ex. 10.17 (3)

自由度 df = n-1=24，查 t 值表可知 : $t_{0.025} = 2.064$

計算 $\bar{X}$ 的臨界值

$\bar{X}_L{}^* = 2.6 - 2.064\dfrac{4.8}{\sqrt{25}} = 0.619$; $\bar{X}_H{}^* = 2.6 + 2.064\dfrac{4.8}{\sqrt{25}} = 4.581$

若 $0.619 \le \bar{X} \le 4.581$ 則接受 H_0

4.計算檢定統計量或P值

t 統計量爲 $t = \dfrac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} = \dfrac{2.4-2.6}{\frac{4.8}{\sqrt{25}}} = -0.208 > t_{24,0.025} = -2.064$

Ex. 10.17 (4)

計算P值

P值 $= 2*P(\bar{X} \le 2.4|\mu=2.6) = 2*P(\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}} \ge \dfrac{2.4-2.6}{4.8/\sqrt{25}})$

$= 2*P(t_{24} \le -0.208) = 2*0.4185 = 0.837 > \alpha = 0.05$

5.下結論

因 $\bar{X}$=2.4 落再接受域且檢定統計量 t=0.208 落在接受域，故不拒絕虛無假設。結論爲：在 α=0.05 下，每人每年的醫療支出約爲 2.6 萬。

Ex. 10.18 (1)

- 有線電視市場調查部門指出，顧客對A、B兩種節目有相同的喜好。現隨機抽取樣本22人，得知喜歡 A 的有58%。問調查部門的報告可靠否？

Ex. 10.18 (2)

題意知 n=225 $\hat{p}$=0.58

1.設立兩假設

若 P=0.05，則市場調查部門的報告可靠，反之則不可靠。

故假設爲 H_0: p=0.5 H_1: $p \ne 0.5$

2.選擇檢定統計量

以 $\hat{p}$ 來檢定，因 p=0.5 q=0.5 。np=225*0.5=112.5

nq=225*0.5=112.5，兩者都大於5，因此是大樣本，

因此採 $(\hat{p}-p_0)/\sigma_p$ 來檢定母體比例。

Ex. 10.18 (3)

3. 選擇顯著水準及決定拒絕域與接受域

選擇 α=0.01，對立假設為 $\neq$，故採雙尾檢定。

故臨界值為2.58及-2.58。

4. 計算檢定統計量

Z檢定統計量為 : Z= $\dfrac{\hat{p}-p_0}{\sigma_p} = \dfrac{0.58-0.5}{\sqrt{\dfrac{0.5*0.5}{225}}} = 2.4$

101

Ex. 10.18 (4)

計算 p 值 = $2\times P(\hat{p}\geq p_0|p=0.5)$

$=2\times P(\dfrac{\hat{p}-p_0}{\sigma_p}\geq \dfrac{0.58-0.5}{\sqrt{0.5*0.5/225}})$

$=2\times P(Z\geq 2.4)=0.0164$

5.下結論

檢定統計量 Z=2.4 小於臨界值 $Z_{0.005}$=2.58，落在接受域

或因 P 值大於 α, 因此拒絕虛無假設，結論為：市場調查部門的報告可靠，兩個節目同樣被喜歡。

102

Exercises

- 1,3,5,7,9,11,15,17,19,21,23,27,29,37,39,41,43, 45,57,61

103