Chapter 10

假設檢定

10.1 假設檢定之概念

統計推論中之假設檢定（testing hypothesis），是對母體中某些參數之可能值爲對或錯之假設或敘述，利用抽樣的樣本資料或實驗結果，再透過機率原理作出拒絕或不拒絕該假設之過程，亦稱之爲顯著性檢定（testing of significance）。

一、基本觀念

在假設檢定中有二種假設要判斷，其中一種假設稱之爲虛無假設（null hypothesis），而另一種則稱爲對立或對抗假設（alternative hypothesis）。茲將其意義定義如下。

定義1

在統計假設檢定中，欲否定的假設稱之為虛無假設，其為一種暫時性的假設，通常以 H_0 表示；而否定虛無假設而被認為對的假設，稱之為對立假設，以 H_1 或 H_a 表示。

一般在定義1中的假設檢定中，又可依其所討論假設中參數的可能數值多寡，而將其分爲簡單假設（simple hypothesis）與複合假設（composite hypothesis）兩大類（見圖10-1）。

定義2

簡單假設是指虛無假設與對立假設中皆只含有一個參數的特定數值的假設；而所謂複合假設，是指虛無假設與對立假設中，參數的假設所含數值並不只一個數值的假設。

例如，$H_0: p = 0.5$；$H_1: p = 0.2$ 或 $H_0: \theta = 10$；$H_1: \theta = 15$ 皆屬於簡單假設，其實簡單假設亦可指爲能完全確定某機率分配之假設。一般而言若一個假設並非簡單假設，即表示其爲複合假設，例如：$H_0: \mu = 8$；$H_1: \mu \neq 8$。

- 假設檢定
  - 簡單假設——$H_0: \theta = \theta_0$；$H_1: \theta = \theta_1$
  - 複合假設
    - 單尾對立假設
      - 右尾對立假設檢定
        - $H_0: \theta = \theta_0$；$H_1: \theta > \theta_0$
        - $H_0: \theta \leq \theta_0$；$H_1: \theta > \theta_0$
      - 左尾對立假設檢定
        - $H_0: \theta = \theta_0$；$H_1: \theta < \theta_0$
        - $H_0: \theta \geq \theta_0$；$H_1: \theta < \theta_0$
    - 雙尾對立假設——$H_0: \theta = \theta_0$；$H_1: \theta \neq \theta_0$

圖10-1　假設檢定之型態

在圖10-1中，可發現單尾對立假設（one-tailed test）又分爲 $H_1: \theta > \theta_0$ 及 $H_1: \theta < \theta_0$ 兩種。通常對立假設 H_1 在右端者（即 $H_1: \theta > \theta_0$），稱之爲右尾對立假設檢定，簡稱爲右尾檢定（right-tailed test）；相反的，當對立假設 H_1 在左端者（即 $H_1: \theta < \theta_0$），稱之爲左尾對立假設檢定，簡稱爲左尾檢定（left-tailed test）；而對立假設 H_1 在雙邊者（即 $H_1: \theta \neq \theta_0$），稱之爲雙尾檢定（two-tailed test）。

定義3

拒絕域（reject region）或稱之為危險域（critical region）是指其為樣本空間中某些數值的集合，而且會導致否定虛無假設的集合，常以 C 或 CR 表示之。

二、兩種類別誤差及風險

㈠誤差的類別：

假設檢定是由抽出的樣本資料去判斷 H_0 及 H_1 兩種假設之真偽，但由於抽樣可能造成樣本偏差，而導致作出錯誤決策，此錯誤之決策有兩種，分別稱之爲型Ⅰ誤差（type I error）及型Ⅱ誤差（type II error），見圖10-2。

真實情況 / 檢定結果	H_0 為真	H_0 不真（H_1 為真）
不拒絕 H_0	正確結論	型Ⅱ誤差
拒絕 H_0	型Ⅰ誤差	正確結論

圖10-2　兩種誤差之觀念

定義4

當虛無假設 H_0 為真時，此檢定的結果卻拒絕虛無假設，稱之為型Ⅰ誤差；而當虛無假設 H_0 不真（或 H_1 為真）時，此檢定之結果卻不拒絕虛無假設，稱之為型Ⅱ誤差。

㈡誤差的機率：

在定義4中所敘述的兩種誤差，其發生機率稱之爲風險，其中型Ⅰ誤差之機率稱爲 α 風險，而型Ⅱ誤差之機率稱爲 β 風險。今將其分別定義如下。

定義5

型Ⅰ誤差發生的機率以 α 表示，稱之為 α 風險，其機率的表示為

$$\alpha = P(拒絕 H_0 \mid H_0 為真)$$
$$= P(落入 C 中 \mid H_0 為真)$$

而型Ⅱ誤差發生的機率以β表示，稱之為β風險，其機率的表示為

$$\beta = P(\text{不拒絕}H_0 \mid H_1\text{為真})$$
$$= P(\text{未落入}C\text{中} \mid H_1\text{為真})$$

定義6

在所有型Ⅰ誤差機率中，發生機率最大者稱為顯著水準（level of significance），又稱之為檢定的大小（size of test），它是在檢定中可容忍的最大型Ⅰ誤差機率，通常以α表示。

茲舉數例來說明定義5之觀念及其機率的計算。

例題 1

在母體為常態分配$N(\mu, \sigma^2)$的假設下，請回答下列有關型一誤差（type I error）與型二誤差（type II error）的問題

單邊檢定（one-sided test）

$H_0 : \mu \le 25$

$H_1 : \mu > 25$

樣本數$n = 81$，標準差$\sigma = 18$，決策規則（decision rule）是：如果$\bar{X} \le 27.56$，接受$H_0(\mu \le 25)$；如果$\bar{X} > 27.56$，拒絕H_0。試問：

(1)若$\mu = 24$，根據上述決策規則，犯型一誤差之機率為何？

(2)若$\mu = 25$，根據上述決策規則，犯型一誤差之機率為何？

(3)若$\mu = 29$，根據上述決策規則，犯型二誤差之機率為何？

（交大管科、元智資管）

Sol：

$$(1)\ \alpha = P(\bar{X} > 27.56 \mid \mu = 24) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{27.56 - 24}{18/\sqrt{81}}\right)$$
$$= P(Z > 1.78) = 0.0375$$

⑵ $\alpha = P(\overline{X} > 27.56 \mid \mu = 25) = P\left(\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \dfrac{27.56-25}{18/\sqrt{81}}\right)$

$= P(Z > 1.28) = 0.1003$

⑶ $\beta = P(\overline{X} \le 27.56 \mid \mu = 29) = P\left(\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \dfrac{27.56-29}{18/\sqrt{81}}\right)$

$= P(Z \le -0.72) = 0.2358$

例題 2

An urn contains 10 chips. An unknown number of the chips are white; the others are red. We wish to test H_0: exactly half the chips are white versus H_1: more than half the chips are white. We draw, without replacement, three chips and reject H_0 if two or more are white .

⑴Find the type I error rate.

⑵Find type II error rate when the urn is 70% white. （98交大管科20%）

Sol：

⑴令θ為白色籌碼個數，則

H_0 ：$\theta = 5$ ；H_1 ：$\theta > 5$

令隨機變數X表示抽出白色個數，且$C = \{X \mid X \ge 2\}$ ，故

$$\alpha = P(\mathrm{I}) = P(X \ge 2 \mid \theta = 5) = \frac{\binom{5}{2}\binom{5}{1}}{\binom{10}{3}} + \frac{\binom{5}{3}\binom{5}{0}}{\binom{10}{3}} = 0.5$$

⑵ $$\beta = P(\mathrm{II}) = P(X < 2 \mid \theta = 7) = \frac{\binom{7}{0}\binom{3}{3}}{\binom{10}{3}} + \frac{\binom{7}{1}\binom{3}{2}}{\binom{10}{3}} = 0.1833$$

例題 3

Suppose the p.d.f. of a random variable x is $f(x)=1/c$, where $0<x<c$. We are to test the null hypothesis $H_0: c=2$ versus the alternative hypothesis $H_1: c=3$. If we draw an observation of x and reject H_0 if the drawn $x>3/5$, then:

⑴What is the probability of committing type Ⅰ error?

⑵What is the probability of committing type Ⅱ error?

（103中央財金12%）

Sol：

⑴因危險域$C=\left\{X \mid X>\frac{3}{5}\right\}$，故型Ⅰ誤差機率為

$$\alpha=P\left(X>\frac{3}{5}\middle| c=2\right)=\int_{\frac{3}{5}}^{2}\frac{1}{2}dx=\left[\frac{x}{2}\right]_{\frac{3}{5}}^{2}=\frac{7}{10}$$

⑵型Ⅱ誤差機率為

$$\beta=P\left(X\le\frac{3}{5}\middle| c=3\right)=\int_{0}^{\frac{3}{5}}\frac{1}{3}dx=\left[\frac{x}{3}\right]_{0}^{\frac{3}{5}}=\frac{1}{5}$$

定義7

在品質管制內容中，一般所謂生產者風險及消費者風險之意義如下：

1. 生產者風險（producer's risk）：即表示生產者所生產的產品多數為良品，但經抽樣檢驗後卻誤判為不良品，因而被拒收的機率，稱之為生產者風險，以α表示。

2. 消費者風險（consumer's risk）：即表示生產者所生產的產品大多數為不良品，但經抽樣檢驗後卻誤判為良品，因而被接受的機率，稱之為消費者風險，以β表示。

例題 4

Suppose we receive a shipment of 2000 items. If our AQL (Acceptable Quality Level) = 0.02, LTPD (Lot Tolerance Percent Defective) = 0.05, and our sampling plan is (sample size $n = 25$, the number of defectives is 1 or less, we accept the shipment). Compute the producer's risk and consumer's risk. （政大企管20%）

Sol：

設隨機變數 X 表示抽中不良品之個數，且 $X \le 1$ 接受該批產品。

⑴生產者風險為

$$\alpha = P(X > 1 \mid p = 0.02) = 1 - \sum_{x=0}^{1} \frac{\binom{40}{x}\binom{1960}{25-x}}{\binom{2000}{25}}$$

$$\doteqdot 1 - \sum_{x=0}^{1} \binom{25}{x} (0.02)^x (0.98)^{25-x} = 0.0886$$

⑵消費者風險為

$$\beta = P(X \le 1 \mid p = 0.05) = \sum_{x=0}^{1} \frac{\binom{100}{x}\binom{1900}{25-x}}{\binom{2000}{25}}$$

$$\doteqdot \sum_{x=0}^{1} \binom{25}{x} (0.05)^x (0.95)^{25-x} = 0.6424$$

例題 5

A book dealer received a lot of 1,500 books from a printer. The dealer plans to sample 12 books and use single-sample acceptance sampling to reach a decision about the lot. If more than one book is defective, the dealer will reject the lot. Suppose the printer is fairly certain that only 3% of the books are defective. What is the producer's risk? Suppose 12% of the lot of books is defective and that this rate would be too high for the dealer to accept. What is the dealer's risk in using this acceptance sampling?　（98中山財管4%）

Sol：

⑴令隨機變數X表示不良品個數，且危險域為$C=\{X\mid X>1\}$，故生產者風險為

$$\alpha=P(X>1\mid p=0.03)=1-\sum_{x=0}^{1}\frac{\binom{45}{x}\binom{1455}{12-x}}{\binom{1500}{12}}$$

$$\doteqdot 1-\sum_{x=0}^{1}\binom{12}{x}(0.03)^x(0.97)^{12-x}\doteqdot 0.0486$$

⑵消費者風險，為

$$\beta=P(X\le 1\mid p=0.12)=\sum_{x=0}^{1}\frac{\binom{180}{x}\binom{1320}{12-x}}{\binom{1500}{12}}$$

$$\doteqdot\sum_{x=0}^{1}\binom{12}{x}(0.12)^x(0.88)^{12-x}\doteqdot 0.5686$$

三、檢定力函數及作業特性函數

㈠檢定力函數：

檢定力函數（power function）及作業特性函數（operating characteristic function）在假設檢定中，分別指拒絕H_0及不拒絕H_0的機率，它可將兩種誤差機率作整合，並藉此二種函數作爲討論假設檢定的基礎，茲將其定義如下：

定義8

檢定虛無假設 H_0 及對立假設 H_1 的檢定力函數，是指抽樣後樣本資料所計算之統計量值落入危險域 C 之機率。換句話說，檢定力函數乃是指在不同參數下，拒絕虛無假設之機率，常以 $PF(\theta)$ 或 $K(\theta)$ 表示，即

$$K(\theta)=P(\text{拒絕}H_0\,|\,\theta)=P(\text{落入}C\text{ 中}|\,\theta)$$

而此檢定力函數之圖形稱為檢定力曲線（power curve）。

Remark

1. 在定義8中，可將檢定力函數寫為

$$K(\theta)=\begin{cases}\alpha(\theta) & \text{，當}\theta\in H_0\\ 1-\beta(\theta) & \text{，當}\theta\in H_1\end{cases}$$

其中 $1-\beta(\theta)$ 稱為在 $\theta\in H_1$ 時，該點下之檢定力（power）。

2. 所謂檢定力是指在 H_1 為真（$\theta_1\in H_1$）下，此檢定之決策規則 C 能查覺虛無假設 H_0 為錯誤之能力，亦即

$$1-\beta=P(\hat{\Theta}\in C\,|\,\theta_1\in H_1)$$

例題 6

茲以例題1為例：

⑴試寫出此檢定之檢定力函數，並繪檢定力曲線。

⑵若樣本數 n 不變，而危險域改為 $C=\left\{\bar{X}\mid\bar{X}>27\right\}$，試問此曲線如何變動（不用計算），請約略繪在同一圖中。

⑶若危險域不變，將 n 變大，試問曲線如何變動？

Sol：

⑴檢定力函數為

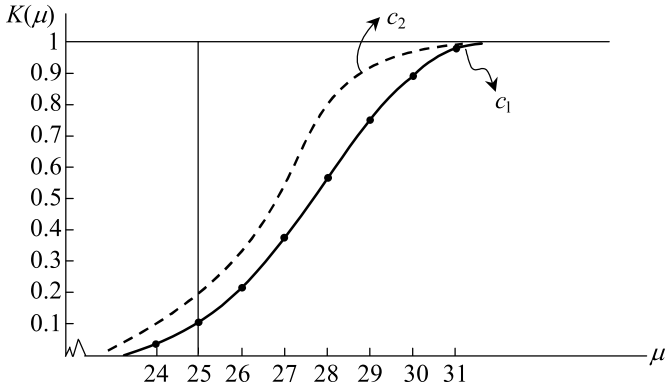
$$K(\mu)=P(\bar{X}>27.56\,|\,\mu)$$

$$=P\left(Z>\frac{27.56-\mu}{18/\sqrt{81}}\right)=P\left(Z>\frac{27.56-\mu}{2}\right)$$

今將 μ 在24, 25, 26, 27, 28, 29, 30, 31各點之值分別計算於下表中

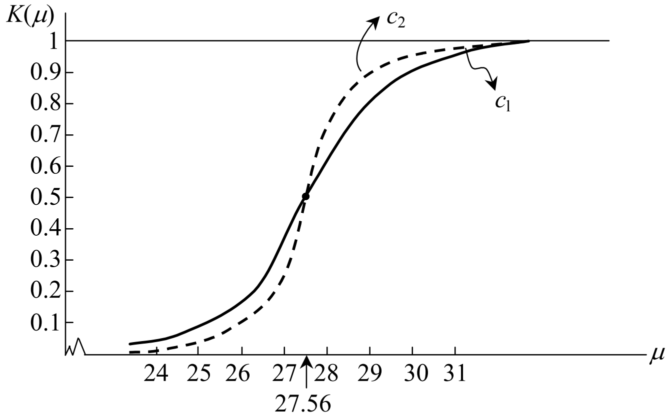
μ	$z = \dfrac{27.56 - \mu}{2}$	$K(\mu) = P(Z > z)$
24	1.78	0.0375
25	1.28	0.1003
26	0.78	0.2177
27	0.28	0.3897
28	–0.22	0.5871
29	–0.72	0.7642
30	–1.22	0.8888
31	–1.72	0.9573

故知檢定力曲線（實線部分）為



⑵若危險域改為 $C = \left\{ \overline{X} \mid \overline{X} > 27 \right\}$，則 $K(\mu)$ 變大，因此檢定力曲線上移至在上圖虛線部分。

⑶若危險域 $C = \left\{ \overline{X} \mid \overline{X} > 27.56 \right\}$ 不變，而樣本數 n 變大，則檢定力曲線為下圖虛線部分。



Remark

由例題6可知下列結論：

1. 在 n 固定下，危險域 C 變大則 α 風險變大，但 β 風險變小，而危險域 C 變小則 α 變小，但 β 變大，且 $\alpha+\beta\neq 1$。

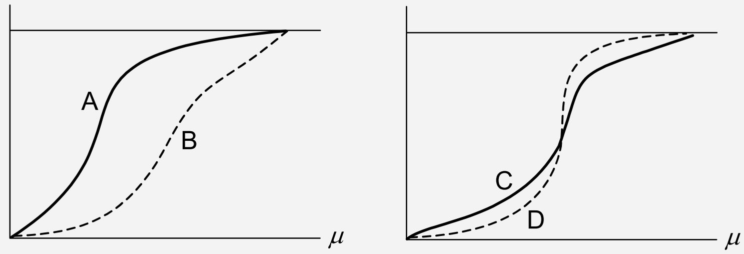
2. 顯著水準發生在虛無假設 H_0 中等號部分。

3. 若要同時減少 α ，β 時，可增加樣本數。

例題 7

In testing $H_0: \mu = 20$ versus $H_a: \mu > 20$ with $\alpha = 0.05$, based on a sample of $n = 30$, the sample mean and standard deviation were given as $\bar{x} = 23$ and $s = 5$.

What would you expect to happen if the power curve moves from curve A to curve B? and what would be the stories if it moves from C to D?



（清大工工）

Sol：

當圖形由A移動到B，則表示危險域的範圍變小。
當圖形由C移動到D，則表示樣本數增加。

㈡作業特性函數：

在定義8中是計算拒絕虛無假設的機率，今若將拒絕的機率改爲不拒絕虛無假設的機率時，則此函數即稱爲作業特性函數，其圖形和檢定力曲線圖形剛好相反，稱之爲作業特性曲線（operating characteristic curve）。

定義9

檢定虛無假設 H_0 及對立假設 H_1 的作業特性函數，是指樣本所計算出來之統計量值未落入危險域 C 之值，亦即在不同 θ 值下，不拒絕 H_0 之機率，常以 $OC(\theta)$ 表示，即

$$OC(\theta)=P(\text{不拒絕}H_0\,|\,\theta)=P(\text{未落入}C\text{ 中}\,|\,\theta)$$

而作業特性函數之圖形稱之為作業特性曲線（OC curve）。

Remark

1. 在定義9中，可將作業特性函數寫為

$$OC(\theta)=\begin{cases}1-\alpha(\theta) & ，\theta\in H_0\\ \beta(\theta) & ，\theta\in H_1\end{cases}$$

2. 在定義8及9之檢定力曲線及作業特性曲線中，若僅將 α 及 β 兩部分畫出之曲線，即稱為錯誤曲線（error curve）。

例題 8

請繪例題6之作業特性曲線及錯誤曲線。

Sol：

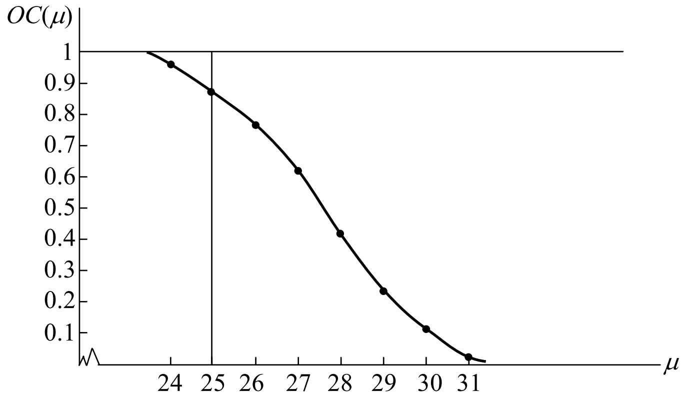
因作業特性函數為

$$\begin{aligned}OC(\mu) &= P(\bar{X} \le 27.56 \mid \mu) \\ &= P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le \frac{27.56-\mu}{18/\sqrt{81}}\right) \\ &= P\left(Z \le \frac{27.56-\mu}{2}\right)\end{aligned}$$

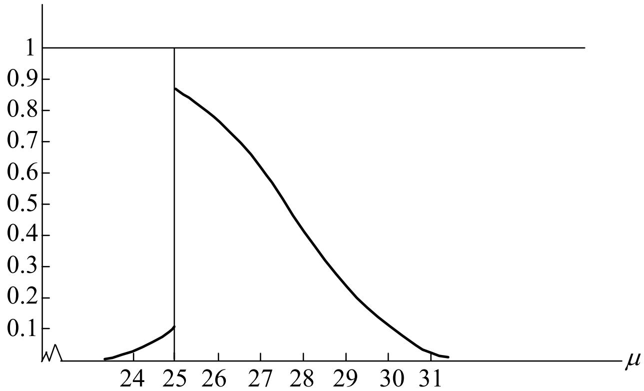
且在不同 μ 之下之機率值計算如下，即

μ	$z = \dfrac{27.56-\mu}{2}$	$OC(\mu) = P(Z \le z)$
24	1.78	0.9625
25	1.28	0.8997
26	0.78	0.7823
27	0.28	0.6103
28	−0.22	0.4129
29	−0.72	0.2358
30	−1.22	0.1112
31	−1.72	0.0427

故知作業特性曲線為



且錯誤曲線為



例題 9

A manufacturing firm evaluates incoming components using the single sampling plan. The single sampling plan has a lot of size $N = 1{,}000$, a sample size $n = 100$, and an acceptance number $c = 2$.
Construct the operating characteristic curve using six points $p = 0.01$, 0.02, 0.03, 0.04, 0.05 and 0.06.　（政大企管25%）

Sol：

令隨機變數 X 表示抽中不良品個數，且 $X \le 2$ 時接受該批產品，因此作業特性函數為

$$OC(p) = P(X \le 2 \mid P = p) = \sum_{x=0}^{2} \frac{\binom{1000p}{x}\binom{1000-1000p}{100-x}}{\binom{1000}{100}}$$

$$\doteqdot \sum_{x=0}^{2} \binom{100}{x} p^x (1-p)^{100-x} \doteqdot \sum_{x=0}^{2} \frac{(100p)^x e^{-100p}}{x!}$$

又在不同參數值 p 之下，其機率分別為

$$OC(0.01) \doteqdot \sum_{x=0}^{2} \frac{1^x e^{-1}}{x!} = 0.920$$

$$OC(0.02) \doteqdot \sum_{x=0}^{2} \frac{2^x e^{-2}}{x!} = 0.677$$

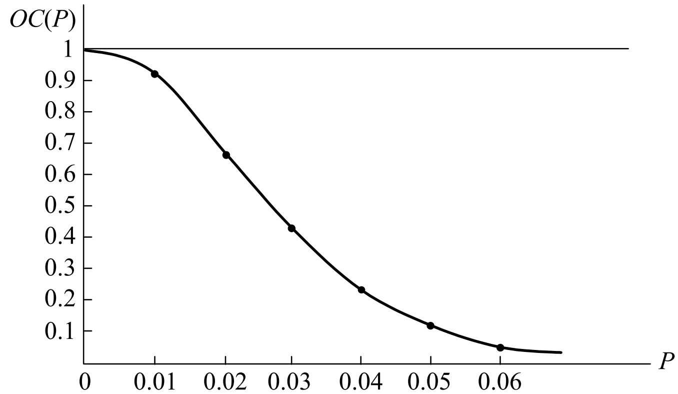
$$OC(0.03) \doteqdot \sum_{x=0}^{2} \frac{3^x e^{-3}}{x!} = 0.423$$

$$OC(0.04) \doteqdot \sum_{x=0}^{2} \frac{4^x e^{-4}}{x!} = 0.238$$

$$OC(0.05) \doteqdot \sum_{x=0}^{2} \frac{5^x e^{-5}}{x!} = 0.125$$

$$OC(0.06) \doteqdot \sum_{x=0}^{2} \frac{6^x e^{-6}}{x!} = 0.062$$

故知作業特性曲線為



四、檢定步驟及檢定之危險域

㈠檢定之步驟：

通常完成一個假設檢定之過程可依下列步驟：

1. 建立虛無假設 H_0 及對立假設 H_1 。

2. 選擇適當之顯著水準 α 。

3. 找出此檢定之統計量及危險域。

4. 利用樣本資料計算統計量之值。

5. 作結論；若 *4.* 之結果落入危險域則拒絕 H_0 ，否則不拒絕 H_0 。

Remark

要如何建立虛無假設及對立假設要看問題而定，通常可將下列之假設放至 H_0：

1. 欲否定之敘述或含有等號之敘述。
2. 在檢定認錯後，將嚴重性較大者放 H_0。

㈡危險域之尋找：

假設檢定中最重要的就是尋找此檢定之統計量與危險域。由於在 n 固定下危險域之大小會影響兩種誤差機率之變動，因此尋找危險域是以控制較嚴重之 α 風險為目標，也就是要尋找一個危險域使得其型 I 誤差發生機率不超過可容忍風險 α，又能使 β 風險達到最小。

Remark

找危險域 C 之原則為先滿足

1. $P(\mathrm{I}) = P(\hat{\Theta} \in C \mid \theta \in H_0) \le \alpha$，再滿足
2. $\min P(\mathrm{II}) = \min P(\hat{\Theta} \notin C \mid \theta \in H_1)$

例題 10

List the five basic procedures in performing a hypothesis test!

（輔大管理10%）

Sol：見上述檢定之五個步驟。

10.2 單一母體平均數 μ 之檢定問題

一、μ 之檢定——以常態分配處理

㈠利用常態分配解決 μ 之檢定問題的條件：

1. 若母體分配未知，但樣本數 $n > 30$ 時，可利用常態分配處理，且若