

第八章

假設與檢定

8.1 前言

在費雪(1935)一書的第二章曾提到下述故事。這故事亦見葉偉文譯(2001)第一章“淑女與下午茶”。

1920年代後期的一個夏日午後，一群大學的研究人員與他們的女眷及訪客，正坐在英國劍橋的戶外桌旁，悠閒地享用下午茶。忽然有位女士宣稱，奶茶的調製順序對風味有很大的影響，把茶加進牛奶裡，和把牛奶加進茶裡，兩者喝起來完全不同。席間那些聰明的科學家，都對這種說法感到可笑，怎麼會不一樣？他們完全無法理解兩種混合結果的化學成分會有什麼差異。但費雪把這當做一個**假設性的問題**，並說“我們來檢定這個命題。”接著設計了一個實驗步驟，包括要準備多少杯茶，依照什麼順序給她喝。

再看民國90年12月20日，Yahoo!奇摩網站上的一則新聞報導。

(中央社記者郭傳信安卡拉十九日專電)據土耳其「時代日報」電子版醫學專欄今天報導，將女性清晨起床後的尿液樣品倒入裝有大麥種子的袋子裡，隔數天後，如果尿液中含有較平常多量的荷爾蒙，大麥就會提前發

芽，即證明女性懷孕。報導中說，古埃及人四千年前即以此法驗孕。

土耳其國立安卡拉大學醫學院婦科系教授庫克在專欄中表示，早在西元前二二〇〇至二〇〇〇年，藥學史極為發達的古埃及人，已能夠不靠化學藥劑即可檢驗出女性是否懷孕。

庫克說，根據至今已發現的古埃及紙草文獻記載，希望知道自己是否懷孕的婦女，必須將自己清晨即起的尿液裝在一個盛有大麥種子的袋子裡，但在此同時，也必須要求另一位確定未懷孕的女性也將清晨即起的尿液裝在另一個有大麥的袋子裡。

庫克表示，由於女性懷孕後，體內會較未懷孕女性產生更多的荷爾蒙，因此泡在懷孕女性尿液中的大麥種子容易發酵並提前發芽，即可確定懷孕，但如果兩袋的大麥種子同時發芽，則證明未懷孕。

庫克最後在文中強調，現代科學已證實這項古老的女性懷孕檢驗法「相當準確」。

在費雪的故事裡，如果只拿一杯奶茶讓那位女士喝，而且她也說對先放茶或先放牛奶，是否會相信她真有能力分辨？可能不會，因有 $1/2$ 的機會她會說對。如果給她兩杯呢？有 $1/4$ 的機會皆說對，不算太小的機率。可能還是不相信她能分辨。如果連續10杯她皆說對，此機率僅為 $1/1,024$ ，算是很小的，即使仍不太相信她有分辨的能力，也許暫時不會排除此一可能性。但是如果20次中錯一次呢？畢竟人難免會犯錯，有時你叫錯朋友的名字，但你不會承認是認錯他。那20次中錯兩次呢？我們對犯錯是有一些忍受的程度，但究竟多大，就因人而異，或因情況而異。附帶一提，如果該女士事先知道10杯中恰5杯先放牛奶，5杯先放茶，則10杯全說對之機率為 $\frac{1}{\binom{10}{5}} = \frac{1}{252}$ 。

再看檢驗女性是否懷孕的方法。俗語說“一樹之果有酸甜之別，一母之子有賢愚之分”。即使同一批大麥，發芽的時間可能便有快慢。就算倒入一沒有懷孕婦女的尿液，大麥可能也會提前發芽。是否要如前述女士喝茶的情況，與10位未懷孕的女性比。“比賽”結果，如果是倒入該婦女尿液的大麥，以6比4提前發芽，那算不算真的領先？要知A、B兩支球隊，即使

勢均力敵(即每場 A, B 獲勝之機率均為 $1/2$)，連比十場， A 隊領先，即至少贏6場的機率為

$$\frac{\binom{10}{6} + \binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}}{2^{10}} = \frac{386}{1,024} \doteq 0.377,$$

此機率並不算低。 B 隊領先的機率也一樣約為0.377。至於平手，即各勝5場的機率則為

$$\frac{\binom{10}{5}}{2^{10}} = \frac{252}{1,024} \doteq 0.246,$$

反而較小。而且怎樣算“提前”？1秒鐘？1分鐘？大約少有人認為這麼小的差異，可算做提前。庫克說現代科學已證實此法相當準確。不過古埃及人是如何操作此法，以得到可靠的結論，就不得而知了。

在一科學試驗裡，並非只有估計母體參數的問題，有時還要依據觀測到的數據，做一些決策。例如，某醫學研究人員，想根據試驗的結果，決定喝咖啡是否會增加致癌的機會。某廠商想決定其產品規格，與標準的規格是否有差異。某社會學家想決定人的血型與眼睛顏色是否獨立。在上述這些情況中，研究人員都是先對某一系統，提出一**猜想**(conjecture)或**假定**(postulate)。然後依據所收集到的數據，而做**決策**。正式地講，在每一情況裡，**猜想**皆可表為**統計假設**(statistical hypothesis, 簡稱**假設**，又hypothesis之複數為hypotheses)的型式。而導致接受或拒絕一統計假設的步驟，就是**統計推論**(statistical inference)之主要工作。所謂統計假設，就是一個關於母體分佈之斷言或猜想。也可以用比較中性的講法：一個關於母體之**敘述**(statement)就是一統計假設。

統計假設與一般的假設有什麼不同？在數學裡我們常有下列這類敘述“假設 $x > y$ ”。由於並未涉及任何隨機的量，所以此非統計假設。但如 μ 表喝咖啡致癌的比例， θ 表不喝咖啡致癌的比例，則“ $\mu > \theta$ ”就是一統計假設了。

除非我們檢視全部母體，否則一統計假設是否為真，常是無法確定的。例如，欲知袋子中紅球的比例。如果袋中球不是特別地多，則全數一遍，便可求出紅球的比例。但若欲知某廠牌電池的平均壽命是否超

過100小時，將全部生產的電池均拿來測試，就不是很實際。欲知高雄市的選民支持候選人甲的比例，是否超過支持候選人乙的比例，也很難訪問每位選民。所以通常的作法是，從我們想了解的母體，取一組隨機樣本，並利用此組樣本，當做是否支持某一假設之證據。如果證據與假設所陳述的假設不吻合，便拒絕該假設，否則便接受該假設。

對於一統計假設，統計學家要設計出一決策步驟。在數學裡，若要證明虛數的單位 $i = \sqrt{-1}$ 與0不能比大小，只要依序證明 $i > 0$ ， $i = 0$ 及 $i < 0$ 皆不成立，便得證了。例如，先假設 $i > 0$ ，則若導出矛盾後，我們便可斷言 $i > 0$ 不成立。但對一隨機現象的假設所做的推論，是可能有錯的。所以在設計決策步驟時，要考慮推論錯誤的機率。要知在處理隨機現象時，統計學家是無法避免犯錯的，但要儘量減小犯錯的機率。要依據所給之限制，提出最佳決策方式。

在數學裡，對於一假設(即一命題)，有成立及不成立兩種結論。有時即使尚未能做結論，數學家知道那只是時機還沒成熟，有如開門的鑰匙仍在找尋中。在統計學裡，我們通常不說一假設成立或不成立。而是說**接受**一假設，或說**拒絕**一假設。拒絕與接受又是什麼意思呢？在數學裡，一個命題如果被接受了(即此命題為真)，就不會有例外地成立。如費馬最後定理(Fermat's last theorem, 見黃文璋(1999a)第四章)：當 $n \geq 3$ 為一整數，且 x, y, z 皆不為0，則 $x^n + y^n = z^n$ 無整數的 x, y, z 解。此定理從被提出至被證出，歷經三百餘年。一旦被證出後，想去找前述方程式的解，便徒勞無功。但在統計學裡，當我們拒絕一假設，並不表該假設為**不可能**(impossible, 英文解釋有not capable of existing or happening, unacceptable等)，而是表該假設“不像會發生”，“似不可信的”(improbable, 英文解釋有unlikely to take place or to be true, doubtful)。

註1.1 口語裡的“不可能”，有時並非真的表不可能，而是指發生的可能性極低。在不可能的任務(Mission Impossible)那部電影裡，主角湯姆克魯斯(Tom Cruise)還是將任務完成了。

某公司宣稱其產品之不良率為 $p = 0.1$ 。此處的假設就是“ $p = 0.1$ ”。隨機地取100個樣本來檢驗，且發現其中有10個不良品，則可能會接受前述假設，因證據並不強到足以推翻 $p = 0.1$ 。但如果將假設改為“ $p = 0.12$ ”，則此假設顯然也不會被推翻。所以，讀者務必要了解的是，接受一假設，僅表示樣本未提供充分的證據以拒絕該假設。另一方面，拒絕一假設，表示樣本提供的證據夠強，足以推翻該假設（但仍有可能犯錯）。以另一方式來說，拒絕一假設，表當該假設為真時，會產生所獲得的樣本之機率很小。例如，對上述情況，若得到20個不良品，則應足以拒絕 $p = 0.1$ 的假設。何以故？若 $p = 0.1$ ，則會得到至少20個不良品之機率為

$$\sum_{i=20}^{100} \binom{100}{i} 0.1^i 0.9^{100-i}。$$

這個機率值雖不太好算，但若利用中央極限定理（且採用連續性的更正），可得近似值

$$\begin{aligned} P(X \geq 20) &= P\left(\frac{X - 10}{\sqrt{100}\sqrt{0.1 \cdot 0.9}} \geq \frac{20 - 10 - 0.5}{\sqrt{100}\sqrt{0.1 \cdot 0.9}}\right) \\ &\doteq 1 - P\left(Z \leq \frac{9.5}{3}\right) \\ &\doteq 1 - 0.9992 = 0.0008, \end{aligned}$$

其中 X 有 $B(100, 0.1)$ 分佈， Z 有 $\mathcal{N}(0, 1)$ 分佈。也就是若 p 真等於0.1，會得到至少20個不良品之機率是很小的，僅約萬分之八。故拒絕“ $p = 0.1$ ”的假設，會犯錯的機率很低。可看出拒絕一假設，表我們認為該假設極可能不真。但接受一假設，有時其他可能性發生的機率卻也不低。因此對一假設，有些人認為以“**不能拒絕**”的說法，取代“**接受**”的說法較適宜。司法院大法官在釋字第485號解釋文中有“其立法意旨與憲法第七條平等原則尚無抵觸。”不說“吻合”而說“尚無抵觸”，與“不能拒絕”一般，都是一種較保守的講法。有點像寧可講“不能說不喜歡”，而不講“喜歡”。不過一般在實際應用時，往往並不那麼謹慎。反正只要理解對於一隨機現象，在這組數據下“接受”，在另一組數據下有可能會拒絕，誤判是難以避免的。因此在文字上那麼在乎，似無必要。

有了一統計假設，下一步就是要去**檢定**此假設是否要接受或拒絕。這整個過程稱為**假設檢定**(testing hypothesis)。假設檢定的理論及架構是波蘭人**奈曼**(Neyman, 1894-1981)及英國人**艾根·皮爾生**(Egon S. Pearson, 1895-1980)，在西元1933年，給出著名的**奈曼-皮爾生引理**(Neyman-Pearson Lemma)所奠定的。在奈曼-皮爾生的架構裡，有一**虛無假設**(null hypothesis)，及一**對立假設**(alternative hypothesis)。虛無假設通常表現況，而對立假設表我們傾向相信的，也就是想證明它為真者。例如，對喝咖啡是否增加致癌機會的問題，若研究人員強烈地認為答案是肯定的，則會將虛無假設取為“喝咖啡不會增加致癌的機會”，而對立假設當然就是取為“喝咖啡會增加致癌的機會”。虛無假設通常以符號 H_0 表之，對立假設則以 H_a (或 H_1, K)表之。

註1.2 奈曼出生於俄國，後來移民到美國，在加州大學柏克萊分校(University of California at Berkeley)任教。近悅遠來，將該校統計系，建立成世界著名的統計研究重鎮。艾根·皮爾生為第六章曾提及的卡爾·皮爾生的獨子，他的出生年代有些書寫1896。奈曼與艾根·皮爾生的討論過程，可參考葉偉文譯(2001)第十及十一章。

統計假設的架構，類似法院審判中的**無罪推定原則**。我國最高法院，於民國二十五年立下的有罪推定原則判例的六十餘年後，終於通過廢止該判例。此後被告原則上是無罪的，不必證明自己無罪，法官只要認為被告罪證不足，即可判無罪，不必窮調查之途，才可以判被告無罪。

我們看民國92年1月14日公佈的修正後之刑事訴訟法第154條第一項：

被告未經審判證明有罪確定前，推定其為無罪。

修正理由中有一句“檢察官須善盡舉證責任，證明被告有罪，俾推翻無罪之推定。”此處的“推定”就是我們用的“假設”。也就是要將“嫌犯無罪”當做虛無假設，檢察官的責任是推翻此一虛無假設(否則也不會起訴該嫌犯了)，但須善盡舉證責任。若非如此，有保護該虛無假設職責的法官，如果認為證據不夠充分，是不會讓虛無假設被推翻的。

無罪推定，可說是較能保障人權，減少冤屈的。以往檢察官若認為法官未窮盡調查之途徑，便判被告無罪，會不服而提起上訴。那是因長期以來法院是採有罪推定原則。但法官判決時，採用無罪推定原則是較有道理的。假設一嫌犯因證據不足而被釋放，如果他是無辜的，那當然很好；如果他其實是有罪，但被釋放後洗面革新，再也不犯罪，那也很好；如果他因心存僥倖或其他原因，又犯了罪，則第二次以後，就不見得每次都有那麼好的運氣了。要知夜路走多總是沒有好下場。

讀者大約也會明白 H_0 為什麼稱為虛無假設了。如果研究人員宣佈喝咖啡不會增加致癌的機會，這種結論可能沒有幾個人有興趣，說了等於白說，很少研究機構公佈這種結果。接受 H_0 ，對研究機構而言，表實驗不算成功。研究人員想要的結論是拒絕 H_0 ！就像檢察官想要的結論是拒絕 H_0 ：嫌犯無罪。此正如一般人有興趣的是電影明星的婚姻有問題(H_a)，當明星的經紀人，一再宣稱該明星夫妻恩愛如常(H_0)，大部分的影迷，對此訊息是不會有興趣的。

在費雪的故事裡，要將 H_0 取成(該女士)“沒有分辨(茶與牛奶何者先放)的能力”， H_a 則取成“有分辨的能力”，推翻 H_0 是大家較有興趣的。而在以大麥檢驗女性是否懷孕的實驗中，當然也是要將 H_0 取成“此法無效”， H_a 取成“此法有效”。

在統計學裡，虛無假設及對立假設，通常表示為關於母體之敘述。而假設檢定則為基於母體之一組隨機樣本，以決定兩個假設中何者該拒絕。底下為一典型的例子。

某城市每日發生之交通事故次數 X ，是以波松分佈為模式。過去參數 λ 一直取為5，自從執行開車不能持手機打電話，且要繫安全帶後，車禍數似乎減少，交通部認為 λ 已減小至3。於是取了幾個月的樣本(車禍數) $X_1, \dots, X_n$ ，想決定 $\lambda = 5$ 或是3。因此可取假設為

$$H_0 : \lambda = 5,$$

$$H_a : \lambda = 3。$$

上述二假設之任一均可決定 X 之分佈，這種假設稱為**簡單假設**(simple hypothesis)。前述二假設 H_0, H_a ，亦有下列寫法

$$H_0 : \lambda = 5, \text{ 對 } H_a : \lambda = 3,$$

或

$$H_0 : \lambda = 5, \text{ vs. } H_a : \lambda = 3,$$

其中vs.為versus(對)之縮寫。曾獲奧斯卡金像獎最佳影片(1979)，由達斯汀霍夫曼(Dustin Hoffman, 他以此片得奧斯卡金像獎最佳男主角)及梅莉史翠普(Meryl Streep, 她以此片得奧斯卡金像獎最佳女配角)所主演的克萊瑪對克萊瑪，英文片名即為Kramer Vs Kramer。

如果是要檢定

$$H_0 : \lambda = 5,$$

$$H_a : \lambda < 5,$$

則 H_0 成為簡單假設，而 H_a 不能完全決定分佈，稱為**複合假設**(composite hypothesis)。在複合假設裡，若為 $\lambda < 5$ 或 $\lambda > 5$ 的型式，便稱為**單邊假設**(one-sided hypothesis)，前者亦稱**下單邊假設**(lower one-sided test)，後者亦稱**上單邊假設**(upper one-sided test)；若為 $\lambda \neq 5$ 的型式，便稱為**雙邊假設**(two-sided hypothesis)。同理也有**單邊檢定**(one-sided test)及**雙邊檢定**(two-sided test)。又“邊”有時稱尾，如**單尾檢定**(one-tailed test)。一般的架構如下：

已知隨機變數 X 的分佈裡有一未知的參數 θ (可以是一向量)， $\theta \in \Omega$ ， Ω 為一我們認為所有可能的 θ 之集合。令 Ω_0, Ω_1 表二 θ 的互斥集合，要檢定

$$H_0 : \theta \in \Omega_0,$$

$$H_a : \theta \in \Omega_1。$$

當 Ω_0 (或 Ω_1)只包含一元素， H_0 (或 H_a)便為簡單假設。又一般 Ω_0 與 Ω_1 不但互斥，且 $\Omega_0 \cup \Omega_1 = \Omega$ ，故常將 Ω_1 寫成 Ω_0^c ，表 Ω_0 之餘集。例如，在檢定

$H_0 : \lambda = 5$, vs. $H_a : \lambda = 3$ 時, Ω 可取成 $\{3, 5\}$, 則 $\Omega_0 = \{5\}$, 且 $\Omega_1 = \{3\} = \Omega_0^c$ 。

註1.3 根據奈曼及皮爾生建立假設檢定理論的原意, 對一假設拒絕或接受, 並非一個對真偽之陳述, 而是用來做為擬採取對策之指引(the rejection and acceptance of the hypothesis were not statements about the truth and falsehood. Those were intended to be the guidelines for future action)。Neyman(1950)一書的pp.259-260 寫著: The terms “accepting” and “rejecting” a statistical hypothesis are very convenient and are well established. It is important, however, to keep their exact meaning in mind and to discard various additional implications which may be suggested by intuition. Thus, to accept a hypothesis H means only to take an action A rather than B . This does not mean that we necessarily believe that the hypothesis H is true. Also, if the application of the rule of inductive behavior “rejects” H , this means only that the rule prescribes action B and does not mean that we believe the H is false.

8.2 基本的假設檢定

根據奈曼-皮爾生的作法, 依據所得到的樣本 $\mathbf{X}$, 基於某一統計量(又稱**檢定統計量**(test statistic)) $T = T(\mathbf{X})$, 來決定是否接受 H_0 。使 H_0 被接受與被拒絕的 T 值之集合, 分別稱為**接受域**(acceptance region)及**拒絕域**(rejection region, 或稱critical region)。前者又稱**允收域**, 後者又稱**拒收域**, 或**棄卻域**。設拒絕 H_0 , 若且唯若 $T(\mathbf{X}) \geq d$, 則拒絕域為 $C = \{T(\mathbf{X}) \geq d\} = \{\mathbf{X} | T(\mathbf{X}) \geq d\}$ 。決定一檢定, 即表決定拒絕域及接受域。由於通常接受域為拒絕域之餘集, 因此所謂**給定一檢定**, 便指給定一拒絕域。我們再定義**檢定函數**(test function)。對一檢定, 以 C 為拒絕域, 即拒絕 H_0 , 若且唯若 $\mathbf{x} \in C$ 。則函數

$$\phi(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in C, \\ 0, & \mathbf{x} \notin C, \end{cases}$$

便稱為**檢定函數**。檢定函數即為拒絕域之指示函數。

對參數 $\theta \in \Omega$, 在檢定 $H_0 : \theta \in \Omega_0$, vs. $H_a : \theta \in \Omega_0^c$ 時, 有兩種犯錯的型式:

1. 當 H_0 為真卻拒絕 H_0 , 此稱**第一型錯誤**(或稱**型 I 錯誤**, type I error)。若 H_0 為簡單假設, 則型 I 錯誤的機率, 通常以符號 α 表之, 即

$$\alpha = P(\text{拒絕}H_0 | H_0 \text{ 為真})。$$

此時 α 也稱為檢定之**顯著水準**(significance level), 又稱**檢定大小**(size of a test, 簡稱size)。若 H_0 為複合假設, 型 I 錯誤的機率, 通常與 $\theta \in \Omega_0$ 有關。此時將顯著水準(簡稱水準)為 α , 取為在這些不同 θ 之下, 型 I 錯誤機率的極大值, 或更一般地說, 取**最小上界**(least upper bound, 即supremum, 簡寫為sup)。即令顯著水準

$$(2.1) \quad \alpha = \sup_{\theta \in \Omega_0} P(\text{拒絕}H_0 | \theta \text{ 為真})。$$

2. 當 H_0 不真卻接受 H_0 , 此稱**第二型錯誤**(或稱**型 II 錯誤**, type II error)。若 H_a 為簡單假設, 則型 II 錯誤的機率, 通常以符號 β 表之, 即

$$\beta = P(\text{接受}H_0 | H_a \text{ 為真})。$$

若 H_a 為複合假設, 則型 II 錯誤的機率, 通常與 $\theta \in \Omega_0^c$ 有關, 可以 $\beta(\theta)$ 表之。即

$$\beta(\theta) = P(\text{接受}H_0 | \theta \text{ 為真}), \theta \in \Omega_0^c。$$

註2.1 設 B 為實數之一非空子集合。若 $K \in R$ 滿足 $x \leq K, \forall x \in B$, 且若 $L < K$, 則存在 $l \in B$, 使得 $l > L$, 則稱 L 為 B 之**最小上界**, 以 $\sup B$ 表之。最小上界若存在必唯一, 但不一定屬於 B 。若 $\sup B \in B$, 此即 B 中

之最大元素。若 $\sup B \notin B$, 則 B 中無最大元素。對 $B = [0, 1]$, 或 $[0, 1)$, $\sup B$ 皆為 1。但前者 $\sup B \in B$, 後者 $\sup B \notin B$ 。同理可定義 B 之最大下界 (greatest lower bound 即 infimum, 簡寫 inf)。

當 H_0 不真且拒絕 H_0 , 這是正確的決定。對 $\forall \theta \in \Omega_0^c$, 此機率, 即 $1 - \beta(\theta)$, 稱為此檢定在 θ 時之檢力 (power)。

理想狀況當然是兩型錯誤的機率皆為 0。但除了少數特例外, 這是不會發生的。一般的情況是, 當樣本數固定時, 兩型錯誤的機率, 有一減小另一必增大。奈曼-皮爾生的作法是: 先給定顯著水準 (通常是一較小的正數), 然後設法求出一有最小型 II 錯誤機率 (或等價地說檢力 $1 - \beta(\theta)$ 最大) 的檢定法。這樣作是基於虛無假設 H_0 應是被保護的。以喝咖啡的例子來說明。如果喝咖啡明明沒事, 卻宣布多喝咖啡易致癌, 那將會引起多大的風波? 又以法院的無罪推定原則為例, 如果一個人其實無罪 (H_0), 卻判他有罪 (型 I 錯誤), 那造成的傷害將難以彌補。所以型 I 錯誤機率的大小, 是應優先控制住的。我們將統計檢定可能的結果列成表 2.1。

表 2.1 統計檢定之可能的結果

	H_0 為真	H_0 不真
接受 H_0	正確	型 II 錯誤
拒絕 H_0	型 I 錯誤	正確

有時為了簡便, 對一事件 A , 可將 $P(A|\theta \text{ 為真})$, 寫成 $P_\theta(A)$ 。又型 I 錯誤的機率及檢力, 皆包含於檢力函數 (power function) $K(\theta)$ 中, 定義為

$$(2.2) \quad K(\theta) = P(\text{拒絕 } H_0 | \theta \text{ 為真}) = P_\theta(\text{拒絕 } H_0), \theta \in \Omega。$$

如果 $\theta \in \Omega_0$, 則 $K(\theta)$ 乃 θ 為真時, 型 I 錯誤的機率, 而顯著水準

$$(2.3) \quad \alpha = \sup_{\theta \in \Omega_0} P_\theta(\text{拒絕 } H_0) = \sup_{\theta \in \Omega_0} K(\theta)。$$

如果 $\theta \in \Omega_0^c$, 則

$$K(\theta) = 1 - \beta(\theta),$$

即 θ 為真時之檢力。

註2.2 在本書裡，我們並不區分顯著水準及檢定大小。但有些書會區分。他們檢定大小的定義與我們相同，仍為 $\sup_{\theta \in \Omega_0} K(\theta) = \alpha$ ；但對 $\forall 0 \leq \alpha \leq 1$ ，顯著水準為 α ，表 $\sup_{\theta \in \Omega_0} K(\theta) \leq \alpha$ 。依其定義，一檢定之顯著水準並不唯一。若顯著水準為 α ，則亦為顯著水準 α_1 ，其中 $0 \leq \alpha_1 \leq \alpha$ 。又在大部分的書中，稱型 I 錯誤及型 II 錯誤為**型 I 誤差**及**型 II 誤差**。不過我們覺得稱做“錯誤”其實較恰當。因我們平常說犯了錯誤，不會說犯了誤差。又在假設檢定裡，符號 α 通常就是指顯著水準，為了簡便，有時我們以“在 α 之下， $\dots$ ”，取代“在顯著水準 α 之下， $\dots$ ”。

例2.1 假設懷疑某一銅板出現正面的機會似較高。令 p 表出現正面的機率，即欲檢定

$$H_0 : p = 0.5,$$

$$H_a : p > 0.5。$$

在此 H_0 為一簡單假設，而 H_a 為單邊假設。投擲銅板 $n = 10$ 次，以出現正面的次數 X 為統計量，則 X 有 $B(10, p)$ 分佈。不難看出拒絕域應取成 X 較大的區域，即 $\{X \geq c\}$ 。對一給定的 α 值，我們想決定拒絕域，即決定 c 之值。對 $p = 0.5, 0.6, 0.7$ ，由附表2得 $P(X \leq k)$ 如下：

$p \setminus k$	0	1	2	3	4	5	6	7	8	9	10
0.5	.001	.011	.055	.172	.377	.623	.828	.945	.989	.999	1
0.6	.000	.002	.012	.055	.166	.367	.618	.833	.954	.994	1
0.7	.000	.000	.002	.011	.047	.150	.350	.617	.851	.972	1

若 $c = 8$ ，即拒絕域為 $\{X \geq 8\}$ ，包含8, 9, 10三個樣本點。則型 I 錯誤之機率(即顯著水準) α ，可由上表第一列($p = 0.5$)獲得：

$$\alpha = P(X \geq 8 | p = 0.5) = 1 - P(X \leq 7 | p = 0.5) \doteq 1 - 0.945 = 0.055。$$

同理，若 $c = 7$ ，則

$$\alpha = P(X \geq 7 | p = 0.5) \doteq 1 - 0.828 = 0.172。$$

反之, 若 $\alpha = 0.055$, 則拒絕域可取為 $\{X \geq 8\}$, 若 $\alpha = 0.172$, 則拒絕域可取為 $\{X \geq 7\}$ 。 α 值愈大, 則拒絕域通常也較大(至少不會更小)。必須一提的是, 對同一 α 值, 有時會有不同的拒絕域。那要選那一個呢? 這點我們稍後再討論。

設 $\alpha = 0.055$, 且取拒絕域為 $\{X \geq 8\}$, 則當 $p = 0.6, 0.7$, 此檢定之型 II 錯誤機率分別為(見上表第二列及第三列)

$$\beta(0.6) = P(X \leq 7 | p = 0.6) \doteq 0.833,$$

$$\beta(0.7) = P(X \leq 7 | p = 0.7) \doteq 0.617。$$

不但隨 p 而變, 且 $\beta(0.7) < \beta(0.6)$, 但二者皆很大。在保護 H_0 , 不輕易拒絕 H_0 (因此 α 要取得小些)之設計下, 使得有時該拒絕 H_0 卻未拒絕, 造成型 II 錯誤之機率不夠小。同樣地, 在無罪推定的原則下, 法庭放走實際有罪的人(型 II 錯誤)之機會, 有時說不定也不低。

設 $\alpha = 0.172$, 且取拒絕域為 $\{X \geq 7\}$ 。則

$$\beta(0.6) = P(X \leq 6 | p = 0.6) \doteq 0.618,$$

且

$$\beta(0.7) = P(X \leq 6 | p = 0.7) \doteq 0.350。$$

印證我們之前指出的: 較大的 α (型 I 錯誤之機率較大), 往往導致 $\beta(\theta)$ 較小(型 II 錯誤之機率較小)。

這裡的機率通常是近似值, 所以以近似值的符號“ $\doteq$ ”表之較恰當。不過有時為了簡便, 常以等號“ $=$ ”表之。

另外, 當拒絕域取為 $\{X \geq 8\}$, 及 $\{X \geq 7\}$ 時, 檢力函數分別為

$$K_1(p) = \sum_{i=8}^{10} \binom{10}{i} p^i (1-p)^{10-i},$$

$$K_2(p) = \sum_{i=7}^{10} \binom{10}{i} p^i (1-p)^{10-i}。$$

可看出當 $p \rightarrow 1$ 時, $K_1(p)$ 及 $K_2(p)$ 皆趨近至 1; 當 $p \rightarrow 0.5$ 時, $K_1(p)$ 及 $K_2(p)$ 分別趨近至其 α 值。

我們再來看, 如果拒絕域取成 $\{X \leq c\}$ 會如何? 這樣的拒絕域當然不合理, 但究竟有什麼後果?

若 $c = 8$, 則 $\alpha = 0.989$, $\beta(0.6) = 0.046$, $\beta(0.7) = 0.149$ 。

若 $c = 7$, 則 $\alpha = 0.945$, $\beta(0.6) = 0.167$, $\beta(0.7) = 0.383$ 。

若 $c = 6$, 則 $\alpha = 0.828$, $\beta(0.6) = 0.382$, $\beta(0.7) = 0.650$ 。

不合理的拒絕域, 導致型 I 錯誤之機率很大, 至於型 II 錯誤之機率則減小。此中原委留給讀者自行討論。

最後我們來看, 那一拒絕域可使型 I 錯誤機率 $\alpha = 0$? 顯然 c 要大於 10, 即 $\alpha = 0$, 若且唯若 $c > 10$ 。但對此最小的 α 值, 對 $\forall \theta > 0.5$, 型 II 錯誤之機率 $\beta(\theta) = 1$, 達到最大值。事實上 α 值為 0, 表永遠接受 H_0 , 因此型 II 錯誤之機率必然為 1。

例2.2 在例2.1中, 改為投擲銅板 $n = 100$ 次, 則出現正面的次數 X 有 $B(100, p)$ 分佈, 雖無法由附表2得到 $P(X \leq k)$, 但可藉助中央極限定理。

如前, 仍取拒絕域為 $\{X \geq c\}$ 。令 $c = 58$ 。則當 $p = 0.5$ 時, $np = 50$, $\sqrt{np(1-p)} = 5$ 。故

$$\begin{aligned}\alpha &= P(X \geq 58 | p = 0.5) \\ &= P\left(\frac{X - 50}{5} \geq \frac{58 - 50 - 0.5}{5} \mid p = 0.5\right) \\ &\doteq P(Z \geq 1.5) \\ &\doteq 1 - 0.9332 = 0.0668,\end{aligned}$$

其中 Z 有 $\mathcal{N}(0, 1)$ 分佈。當 $p = 0.6$ 時, $np = 60$, $\sqrt{np(1-p)} = \sqrt{24}$ 。故型 II 錯誤之機率

$$\begin{aligned}\beta(0.6) &= P(X \leq 57 | p = 0.6) \\ &= P\left(\frac{X - 60}{\sqrt{24}} \leq \frac{57 - 60 + 0.5}{\sqrt{24}} \mid p = 0.6\right) \\ &\doteq P(Z \leq -0.5103)\end{aligned}$$

$$\doteq 1 - 0.6951 = 0.3049。$$

與例2.1中 $n = 10$, 且差不多大小之 α 值(即0.055)相比, 當 $p = 0.6$ 時型II錯誤的機率(0.833), 顯然下降很多。這是一般的情況, 在 α 值不變下, 增加樣本數, 可減小型II錯誤的機率。

即使樣本的分佈不知, 但變異數知道, 則只要樣本數夠多(譬如說至少有30), 通常便可藉助中央極限定理, 以建立一檢定, 見下例。

例2.3 假設某廠牌罐裝咖啡的標示容量為90(公克), 雖以機器裝罐, 但仍有誤差, 由過去資料知標準差為2(公克)。由於咖啡若容量不足, 顧客會抱怨; 容量過多, 罐子鼓脹, 且增加成本, 故容量不足或過多均不很恰當。以 μ (公克)表所裝咖啡量之期望值, 我們想檢定 μ 是否為90。即要檢定

$$H_0 : \mu = 90,$$

$$H_a : \mu \neq 90。$$

取 n 罐咖啡來檢驗, 以 $\bar{X}$ 表容量之樣本平均。由中央極限定理, $\bar{X}$ 有近似的 $\mathcal{N}(\mu, \sigma/\sqrt{n})$ 分佈。在型I錯誤機率 $\alpha = 0.05$ 之下, 取拒絕域為 $\{|Z_0| > c\}$, 其中

$$Z_0 = \frac{\bar{X} - \mu}{2/\sqrt{n}}。$$

c 之值可由在 H_0 之下, 即 $\mu = 90$, 而

$$P\left(\left|\frac{\bar{X} - 90}{2/\sqrt{n}}\right| > c\right) = 0.05$$

來決定。由附表4, 得 $c \doteq 1.96$ 。

現假設 $n = 30$, 且 $\bar{X} = 91.3$, 則

$$|Z_0| = \left|\frac{91.3 - 90}{2/\sqrt{30}}\right| \doteq 3.56 > 1.96,$$

故在顯著水準 $\alpha = 0.05$ 之下, 拒絕 $H_0 : \mu = 90$, 即認為咖啡容量的期望值並非90(公克)。事實上證據顯示咖啡容量的期望值大於90(公克)。

在上例中，爲了符號上的簡便，對於 n 個樣本之平均，以 $\bar{X}$ 表之，而不如以往用 $\bar{X}_n$ 表之。一般而言，在母體變異數 σ^2 已知下，欲對母體之期望值 μ 做一些推論。先由該母體產生一組隨機樣本 $X_1, \dots, X_n$ 。由前兩章的討論，我們知道樣本平均 $\bar{X}$ 常用來當做 μ 之估計量。如果母體有常態分佈，則 $(\bar{X} - \mu)/(\sigma/\sqrt{n})$ 有 $\mathcal{N}(0, 1)$ 分佈。如果母體並非常態分佈，但樣本數 n 夠大，則

$$(2.4) \quad Z_0 = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}$$

有近似的常態分佈。取 Z_0 爲檢定統計量，且對 $\forall 0 < p < 1$ ，令 z_p 滿足

$$(2.5) \quad P(Z \leq z_p) = p,$$

其中 Z 有 $\mathcal{N}(0, 1)$ 分佈， z_p 即爲 $\mathcal{N}(0, 1)$ 分佈之第 p 分位數。若欲檢定的 H_0 爲簡單假設， H_a 爲雙邊假設，即

$$H_0 : \mu = \mu_0,$$

$$H_a : \mu \neq \mu_0,$$

則在型 I 錯誤機率爲 α 之下，拒絕域(或是近似的拒絕域，底下不論母體是否爲常態分佈，皆稱拒絕域，而省略“近似的”三字)可取爲

$$(2.6) \quad \begin{aligned} &\{|Z_0| > z_{1-\alpha/2}\} \\ &= \{\bar{X} > \mu_0 + z_{1-\alpha/2}\sigma/\sqrt{n}, \text{ 或 } \bar{X} < \mu_0 - z_{1-\alpha/2}\sigma/\sqrt{n}\}. \end{aligned}$$

此因在 H_0 之下， $P(|Z_0| > z_{1-\alpha/2}) = \alpha$ 。圖2.1給出拒絕域及接受域。

若欲檢定的 H_0 爲簡單假設， H_a 爲上單邊假設，即

$$H_0 : \mu = \mu_0,$$

$$H_a : \mu > \mu_0,$$

則在型 I 錯誤機率爲 α 之下，拒絕域可取爲

$$(2.7) \quad \{Z_0 > z_{1-\alpha}\} = \{\bar{X} > \mu_0 + z_{1-\alpha}\sigma/\sqrt{n}\}.$$

圖2.2給出拒絕域及接受域。

若欲檢定的 H_0 為簡單假設, H_a 為下單邊假設, 即

$$H_0 : \mu = \mu_0,$$

$$H_a : \mu < \mu_0,$$

則在型 I 錯誤機率為 α 之下, 拒絕域可取為

$$(2.8) \quad \{Z_0 < -z_{1-\alpha}\} = \{\bar{X} < \mu_0 - z_{1-\alpha}\sigma/\sqrt{n}\}。$$

圖2.3給出拒絕域及接受域。

長久以來型 I 錯誤機率(或顯著水準) α 之值, 常取成0.1, 0.05或0.01等。若 H_a 為單邊假設, 則對應 $z_{0.9} \doteq 1.282$, $z_{0.95} \doteq 1.645$, $z_{0.99} \doteq 2.326$ 。若 H_a 為雙邊假設, 則對應 $z_{0.95} \doteq 1.645$, $z_{0.975} \doteq 1.96$, $z_{0.995} \doteq 2.576$ 。

若母體變異數 σ^2 不知(但假設 σ^2 有限), 如何得到一個對母體期望值 μ 之檢定? 先設母體有常態分佈, 且 $X_1, \dots, X_n$ 為由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本, 則由第四章例3.2,

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

有 $\mathcal{T}_{n-1}$ 分佈, 其中 S 為樣本標準差。現對 $\forall 0 < p < 1$, $m \geq 1$, 令 $t_{p,m}$ 為 $\mathcal{T}_m$ 分佈之第 p 分位數。即若 T 有 $\mathcal{T}_m$ 分佈, 則

$$(2.9) \quad P(T \leq t_{p,m}) = p。$$

在顯著水準為 α 之下, 若欲檢定 $H_0 : \mu = \mu_0$, vs. $H_a : \mu \neq \mu_0$, 則拒絕域可取為

$$\left\{ \left| \frac{\sqrt{n}(\bar{X} - \mu_0)}{S} \right| > t_{1-\alpha/2, n-1} \right\};$$

若欲檢定 $H_0 : \mu = \mu_0$, vs. $H_a : \mu > \mu_0$, 則拒絕域可取為

$$\left\{ \frac{\sqrt{n}(\bar{X} - \mu_0)}{S} > t_{1-\alpha, n-1} \right\};$$

若欲檢定 $H_0 : \mu = \mu_0$, vs. $H_a : \mu < \mu_0$, 則拒絕域可取為

$$\left\{ \frac{\sqrt{n}(\bar{X} - \mu_0)}{S} < -t_{1-\alpha, n-1} \right\}。$$

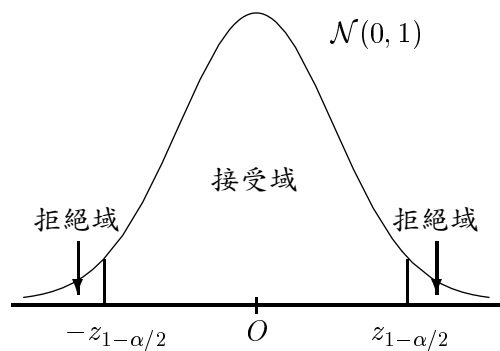


圖2.1 檢定 $H_0 : \mu = \mu_0$, vs. $H_a : \mu \neq \mu_0$

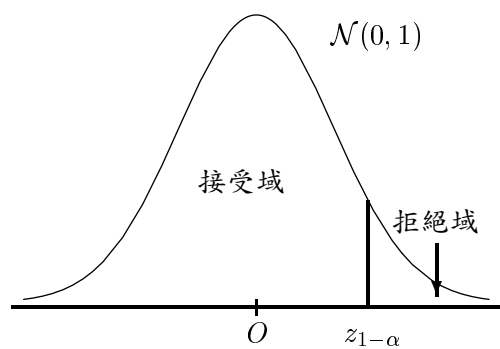


圖2.2 檢定 $H_0 : \mu = \mu_0$, vs. $H_a : \mu > \mu_0$

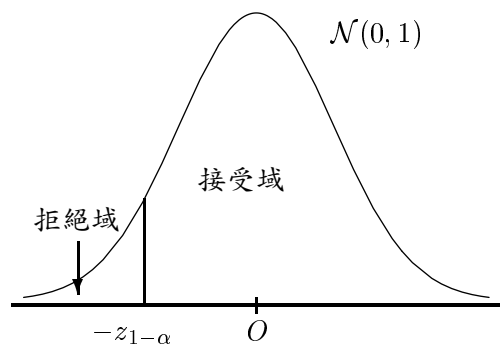


圖2.3 檢定 $H_0 : \mu = \mu_0$, vs. $H_a : \mu < \mu_0$

這種檢定稱為單樣本之 t 檢定(one-sample t -test)。

例2.4 假設某公司宣稱所生產的電池平均壽命為46小時。隨機地取12個電池測試, 得平均使用 $\bar{X} = 42$ (小時), 樣本標準差 $S = 11.9$ (小時)。假設電池壽命有常態分佈, 試在 $\alpha = 0.05$ 之下, 判定電池平均壽命是否小於46小時。

解.依題意, 欲檢定 $H_0: \mu = 46$ (小時), vs. $H_a: \mu < 46$ (小時)。由附表6得 $-t_{0.95,11} \doteq -1.796$ 。因此拒絕域為

$$\left\{ \frac{\sqrt{12}(\bar{X} - 46)}{S} < -1.796 \right\}。$$

現因

$$\frac{\sqrt{12}(42 - 46)}{11.9} \doteq -1.164 > -1.796。$$

故不能拒絕 H_0 。由於 $-t_{0.9,11} \doteq -1.363 < -1.164$, 故即使讓 α 大些, 如取 $\alpha = 0.1$, 仍不能拒絕 H_0 。若取 $\alpha = 0.15$, 則 $-t_{0.85,11} = -1.088 > -1.164$, 便可拒絕 H_0 。不過這樣的 α 太大了些。

雖然樣本平均42明顯地比欲檢定的值46小, 但依據所得的資料, 我們只能說公司所宣稱的 $\mu = 46$, 尚無不合理。事實上在 $\alpha = 0.05$ 之下, 即使得到樣本平均40小時, 且樣本標準差仍維持11.9, 仍無法拒絕 H_0 :

$$\frac{\sqrt{12}(40 - 46)}{11.9} \doteq -1.747 > -1.789。$$

讀者可能已看出關鍵在何處了, 就是所得樣本之標準差太大! 若樣本標準差減半為5.95, 樣本平均仍設為42, 則即使 α 小些, 如取 $\alpha = 0.025$, 都會拒絕 H_0 :

$$\frac{\sqrt{12}(42 - 46)}{5.95} \doteq -2.329 < -2.201 = -t_{0.975,11}。$$

標準差太大, 表樣本會有較大的變異, 因此即使真正的平均為46, 得到一離它太遠的樣本平均, 是很可能發生的。

在很多實際的情況中, 不但母體不一定有常態分佈, 變異數通常也不知道。但只要樣本數夠多(如至少有30), 且母體變異數存在, 則中央極

限定理仍適用，且可以 S 取代 σ (見第四章例2.14)。即 $\sqrt{n}(\bar{X} - \mu)/S$ 有近似的 $\mathcal{N}(0, 1)$ 分佈。如此仍可對母體之期望值 μ 建立一檢定。當然這是大樣本下的結果。底下給一例子。

例2.5 假設欲檢定某廠牌之輪胎，是否如其所宣稱的平均可行駛22,000公里。令 $H_0 : \mu = 22,000$, vs. $H_a : \mu < 22,000$ 。隨機地取100個輪胎測試，得平均行駛21,431公里，樣本標準差為1,295公里。取 $\alpha = 0.01$ ，得 $z_{0.99} \doteq 2.326$ 。因

$$\frac{\sqrt{100}(21,431 - 22,000)}{1,295} \doteq -4.394 < -2.326,$$

故拒絕 H_0 。換句話說在 $\alpha = 0.01$ 之下，判定該廠牌之輪胎，並不如其所宣稱的那麼好。讀者可看出，即使取 $\alpha = 0.001$ ，亦會拒絕 H_0 。

在上例中，由於樣本數夠多，故雖母體分佈不知，且變異數亦不知，仍可建立一檢定。基本上，對檢定母體之期望值，且虛無假設為簡單假設，則只要母體有常態分佈或樣本數夠多，檢定便不是問題。其他的情況，就要視母體的分佈而定。例如，若母體有伯努力分佈，則仍可檢定其期望值(見例2.1)。

檢力函數在檢定時有重要的角色，在8.5節我們會闡釋。此處先以兩個較簡單的例子，來討論檢力函數。

例2.6 設 X 有 $B(5, \theta)$ 分佈，欲檢定 $H_0 : \theta \leq 0.5$, vs. $H_a : \theta > 0.5$ 。以 X 為檢定統計量。先考慮拒絕域為 $\{X = 5\}$ 。則檢力函數

$$K_1(\theta) = P_\theta(X = 5) = \theta^5。$$

圖2.4繪出 $K_1(\theta)$ 之圖形。易見若取上述拒絕域，則

$$\text{型 I 錯誤之機率} \leq \sup_{\theta \in [0, 0.5]} K_1(\theta) = K_1(0.5) = 0.5^5 = 0.03125 = \alpha_1,$$

其中 α_1 表顯著水準。但當 $\theta > 0.5$ 時，對大部分的 θ ，型 II 錯誤之機率 $\beta_1(\theta) = 1 - K_1(\theta)$ 均太大。只有當 θ 較接近1時， $\beta_1(\theta)$ 才會較小。例如，欲使 $\beta_1(\theta) < 0.5$ ，則

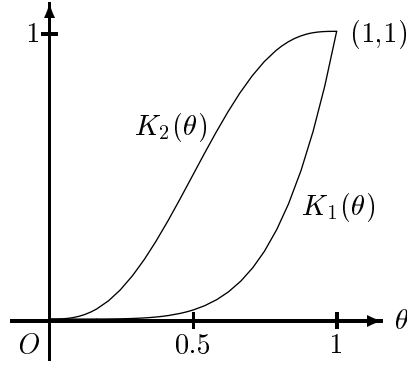


圖2.4 例2.6中之檢力函數

$$K_1(\theta) = \theta^5 > 0.5 \Leftrightarrow \theta > 0.5^{1/5} \doteq 0.87;$$

欲使 $\beta_1(\theta) < 0.1$, 則

$$K_1(\theta) = \theta^5 > 0.9 \Leftrightarrow \theta > 0.9^{1/5} \doteq 0.979。$$

其次考慮拒絕域為 $\{X = 3, 4, 5\}$, 顯然型 II 錯誤之機率會小些。此時檢力函數為

$$K_2(\theta) = P_\theta(X = 3, 4, 5) = 10\theta^3(1-\theta)^2 + 5\theta^4(1-\theta) + \theta^5。$$

$K_2(\theta)$ 之圖形亦繪在圖2.4中。由 $K_2(\theta)$ 之圖形皆在 $K_1(\theta)$ 圖形之上方, 顯示第二種檢定型 I 錯誤之機率較大, 但型 II 錯誤之機率 $\beta_2(\theta) = 1 - K_2(\theta)$ 較小。又 $\theta \rightarrow 0$ 時, $K_1(\theta)$ 及 $K_2(\theta)$ 均趨近至 0; $\theta \rightarrow 1$ 時, $K_1(\theta)$ 及 $K_2(\theta)$ 均趨近至 1。

讀者是否可看出, 若取拒絕域為 $\{X = 4, 5\}$, 則其檢力函數 $K_3(\theta)$ 之圖形介於 $K_1(\theta)$ 與 $K_2(\theta)$ 之間。

例2.7 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本, σ 設為已知。欲檢定 $H_0: \mu \leq \mu_0$, vs. $H_a: \mu > \mu_0$ 。取 $\bar{X}$ 為檢定統計量, 且取拒絕域為 $\{\bar{X} > c\}$ 。則檢力函數

$$\begin{aligned}
 (2.10) \quad K(\mu) &= P_{\mu}(\bar{X} > c) = P_{\mu}\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{c - \mu}{\sigma/\sqrt{n}}\right) \\
 &= P\left(Z > \frac{c - \mu}{\sigma/\sqrt{n}}\right) = 1 - \Phi\left(\frac{c - \mu}{\sigma/\sqrt{n}}\right),
 \end{aligned}$$

其中 Z 有 $\mathcal{N}(0, 1)$ 分佈。圖 2.5 給出 $y = K(\mu)$ 之圖形(習題第 11 題)。 $K(\mu)$ 爲 $-\mu$ 之嚴格漸增函數, 且滿足

$$\lim_{\mu \rightarrow -\infty} K(\mu) = 0, \quad \lim_{\mu \rightarrow \infty} K(\mu) = 1。$$

若設定顯著水準爲 α , 則

$$(2.11) \quad \alpha = \sup_{\mu \leq \mu_0} K(\mu) = K(\mu_0) = 1 - \Phi\left(\frac{c - \mu_0}{\sigma/\sqrt{n}}\right)。$$

由上式解出

$$(2.12) \quad c = \mu_0 + z_{1-\alpha}\sigma/\sqrt{n}。$$

由於通常 $\alpha < 0.5$, 故 $z_{1-\alpha} > 0$, 且 $c > \mu_0$ 。

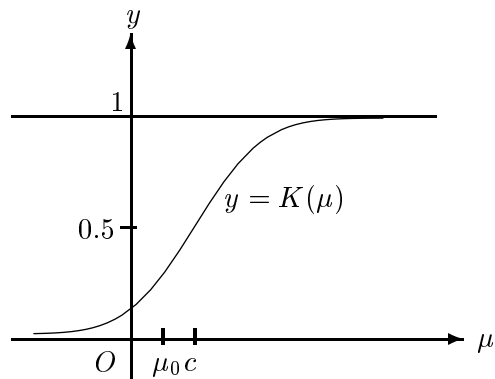


圖 2.5 例 2.7 中之檢力函數

現設 $n = 25$, $\sigma = 2$, $\mu_0 = 0$, 且取 $\alpha = 0.05$ 。則由 (2.12) 式得 $c = 0.4z_{0.95} = 0.4 \cdot 1.645 = 0.658$ 。即此時, 拒絕域爲 $\{\bar{X} > 0.658\}$ 。

由於型II錯誤機率

$$\beta(\mu) = 1 - K(\mu), \mu > \mu_0,$$

爲一嚴格漸減函數, 因此

$$\sup_{\mu > \mu_0} \beta(\mu) = \beta(\mu_0) = 1 - K(\mu_0) = 1 - \alpha,$$

一般而言是較大的(因 α 通常較小)。我們無法要求型II錯誤機率對每一 μ 皆很小。但可以要求 μ 大於某一值之後, $\beta(\mu)$ 均不超過某一定值。例如, 取 $\alpha = 0.1$, 且當 $\mu \geq \mu_0 + \sigma$ 時(即 μ 大於 μ_0 至少一個標準差), 要求型II錯誤之機率不超過0.2。此時該如何選取 c (表示成 μ_0 及 σ 之函數)及樣本數 n ?

如前, 由 $\alpha = 0.1$, 先得到

$$K(\mu_0) = 1 - \Phi\left(\frac{c - \mu_0}{\sigma/\sqrt{n}}\right) = 0.1,$$

再令

$$\beta(\mu_0 + \sigma) = 1 - K(\mu_0 + \sigma) = \Phi\left(\frac{c - (\mu_0 + \sigma)}{\sigma/\sqrt{n}}\right) = 0.2。$$

利用 $z_{0.9} \doteq 1.282$, $z_{0.2} = -z_{0.8} \doteq -0.842$, 得

$$c - \mu_0 = 1.282\sigma/\sqrt{n},$$

$$c - \mu_0 - \sigma = -0.842\sigma/\sqrt{n}。$$

解出 $n \doteq 2.124^2 \doteq 4.511$ 。因 n 須爲正整數, 且較大的 n , 可使 $\beta(\mu)$ 較小, $\mu > \mu_0$, 故 n 至少要取成5。當 $n = 5$, $c = \mu_0 + 1.282\sigma/\sqrt{5}$, 此時 $\mu \geq \mu_0 + \sigma$ 時, $\beta(\mu)$ 均不超過

$$\begin{aligned} \beta(\mu_0 + \sigma) &= \Phi\left(\frac{c - (\mu_0 + \sigma)}{\sigma/\sqrt{n}}\right) \doteq \Phi\left(\frac{\mu_0 + 1.282\sigma/\sqrt{5} - (\mu_0 + \sigma)}{\sigma/\sqrt{5}}\right) \\ &\doteq \Phi(-0.9541) = 1 - \Phi(0.9541) \doteq 1 - 0.83 = 0.17。 \end{aligned}$$

但型II錯誤機率之最小上界爲 $\beta(\mu_0) = 1 - K(\mu_0) = 1 - 0.1 = 0.9$, 算是相當大。

顯著水準 α 取成0.1, 0.05或0.01等, 只是一種習慣。如果假設 H_0 爲真, 但卻有一小機率的事件發生(即檢定統計量落在拒絕域, 見圖2.1-2.3), 我

們自然可以“合理地懷疑”原先的假設 H_0 不真。當然 $\alpha = 0.05$ 或 0.01 是否已經夠小了，要視情況而定。如果是 H_0 ：曾參沒殺人，則 α 值很可能會取得比 0.01 小很多才行。否則若錯誤地推翻 H_0 ，判曾參重刑，其後果便很難彌補。至於對產品規格的品質， α 取成 0.01 就已夠小了。此因即使“誤怪”廠商，廠商往往只有一些金錢上的損失，尚可容忍。

例2.8 某賭場有一以銅板為賭具的賭局。某人懷疑賭場的銅板並非公正。經投擲100次後，出現63次正面。當銅板為公正時，出現正面數 X 至少是63次之機率為

$$\begin{aligned} P(X \geq 63) &= P\left(\frac{X - 50}{\sqrt{100 \cdot 0.5 \cdot 0.5}} \geq \frac{63 - 50 - 0.5}{\sqrt{100 \cdot 0.5 \cdot 0.5}}\right) \\ &\doteq P(Z \geq 2.5) \doteq 1 - 0.9938 = 0.0062. \end{aligned}$$

其中 Z 有 $\mathcal{N}(0, 1)$ 分佈。此機率值很小，賭徒當然懷疑賭場作假。

由於 $63 - 50 = 13$ ，若考慮出現正面數 X ，與公正銅板出現正面數的期望值50，相差至少是13，則機率為

$$P(|X - 50| \geq 13) \doteq 2P(Z \geq 2.5) \doteq 0.0124,$$

仍是很小。

另外，如果出現65次正面，則因

$$P(X \geq 65) \doteq P(Z \geq 2.9) \doteq 1 - 0.9981 \doteq 0.0019.$$

雖只是多兩次正面，機率值卻小很多，更加強懷疑了。

不過要注意，無論出現再多的正面，我們只能“懷疑”銅板之公正性，畢竟機率為正的事件總是可能發生的（並非一不可能的事件）。在黃文璋(1999b)一文中的“當稀有碰到大數”那一節，我們已做了一些說明。在本例中，雖 $P(X \geq 65) \doteq 0.0019$ （千分之1.9）很小，但若觀察1,000回，每回投擲銅板100次，則其中出現一、兩回65次以上的正面，並不足為奇。此時若懷疑賭場之公正性，就有點大驚小怪了。

又如某大公司，有一萬名員工，在年終模彩時，總經理的秘書中了首獎。有人懷疑其中有作弊，並作一檢定如下：

$$H_0 : \text{沒有作弊, vs. } H_a : \text{有作弊。}$$

因在 H_0 之下，隨機地抽獎，總經理秘書會中首獎之機率為萬分之一(0.0001)，故若取拒絕域為“總經理秘書中首獎”，則型 I 錯誤之機率為0.0001。因此當總經理秘書中首獎，在絕大部分常見的顯著水準之下，會拒絕 H_0 而接受 H_a 。但這樣的結論合理嗎？用類似的討論，可得到全公司每一員工都不該中首獎！

例2.8便引進**p-值**(*p-value*)的概念。

在假設檢定裡，常先定下一顯著水準 α 值，而做出拒絕或接受虛無假設 H_0 的決定。但有時可能會對這樣表示結論的方式不盡滿意。因即使宣佈接受 H_0 ，並未指出究竟是勉強接受 H_0 (檢定統計量雖落在接受域，但接近邊緣)，或很安心地接受 H_0 。比如說若僅是宣佈某位學生通過測驗，提供的資訊當然不夠多。

為避免上述困擾，在實務裡，便常有求**p-值**的作法。所謂**p-值**，乃在 H_0 為真之下，檢定統計量會落在比觀測值至少同樣**極端**(extreme)的區域之機率。求出**p-值**之後，不同的決策者，可依其所設定的 α 值，而做出決定。對一觀測值，**p-值**亦可等價地定義為要拒絕 H_0 之最小的顯著水準之值。因此**p-值**又稱**觀測大小**(observed size)。設 Z_0 之定義如(2.4)式。則當 $Z_0 = z$ ，**p-值**為

$$p = \begin{cases} 2(1 - \Phi(|z|)) & , \text{檢定: } H_0 : \mu = \mu_0, \text{ vs. } H_a : \mu \neq \mu_0, \\ 1 - \Phi(z) & , \text{檢定: } H_0 : \mu = \mu_0, \text{ vs. } H_a : \mu > \mu_0, \\ \Phi(z) & , \text{檢定: } H_0 : \mu = \mu_0, \text{ vs. } H_a : \mu < \mu_0. \end{cases}$$

求出**p-值**後，若 $p < \alpha$ ，則拒絕 H_0 ，否則接受 H_0 。

同一觀測值及同一 H_0 ，**p-值**乃與對立假設有關係。

“顯著”的英文significant之意義為1. Having or expressing a meaning; meaningful, 2. Notable等。如果所得觀測值的**p-值**夠小，小於 α ，則此組觀

測值導致棄卻虛無假設，因此這組觀測值是有意義的，也就是其效果是顯著的，或說該結果具**顯著性**。顯著水準也因此又稱**有意水準**，指一有意義的水準。

例2.9 某人投擲標槍100次，得平均距離71.8公尺，過去資料顯示投擲距離之標準差為8.9公尺。試問是否可宣佈平均投擲距離超過70公尺。

解.依題意，要檢定

$$H_0 : \mu = 70,$$

$$H_a : \mu > 70。$$

因樣本數夠多(有100)，故可以常態分佈來近似。先求出

$$Z_0 = \frac{71.8 - 70}{8.9/\sqrt{100}} \doteq 2.02,$$

又

$$P(Z > 2.02) \doteq 1 - 0.9783 = 0.0217。$$

即得p-值約為0.0217。故若設定顯著水準 $\alpha = 0.05$ ，則拒絕 H_0 ；若設定 $\alpha = 0.01$ ，則不能拒絕 H_0 。

由圖2.1-2.3可看出，對 H_0 為簡單假設，在同一顯著水準之下，即使得到相同的觀測值，有可能 H_a 為雙邊假設時是接受 H_0 ，而 H_a 為單邊假設時是拒絕 H_0 。其中原理應不難想通。底下為一例子。

例2.10 某地西元1962年時，男子平均身高為171公分，標準差為7公分。在西元2002年，隨機地抽取40名男子，得平均身高173公分。試在 $\alpha = 0.05$ 之下，分別檢定該地男子(i)平均身高改變，(ii) 平均身高增加。

解.以 μ 表平均身高，則 $H_0 : \mu = 171$ ，且在(i)中 $H_a : \mu \neq 171$ ，在(ii)中 $H_a : \mu > 171$ 。仿例2.9之作法，得

$$Z_0 = \frac{173 - 171}{7/\sqrt{40}} \doteq 1.807。$$

對檢定(i), p -值約為 $2(1 - \Phi(1.807)) \doteq 0.0707$, 對檢定(ii), p -值約為 $1 - \Phi(1.807) \doteq 0.0353$ 。故在 $\alpha = 0.05$ 之下, 接受(i)中之 H_0 , 即該地男子平均身高沒有改變, 拒絕(ii)中之 H_0 , 即接受該地男子平均身高增加。

在本節一開始(亦見例2.1), 我們曾說當樣本數固定, 兩型錯誤的機率, 通常有一減小另一必增大。但有時並非如此, 尤其是檢定統計量不同時, 見下例。

例2.11 設 $X_1, \dots, X_n$ 為一組由 $U[0, \theta]$ 分佈所產生之隨機樣本, $\theta \in \Omega = (0, \infty)$ 。欲檢定 $H_0: 5 \leq \theta \leq 6$, vs. $H_a: \theta < 5$ 或 $\theta > 6$ 。

首先 $X_{(n)}$ 為 θ 之一充分統計量。我們先考慮基於 $X_{(n)}$ 之檢定。 $X_{(n)} > 6$ 時, 當然要拒絕 H_0 , 因在 H_0 之下, $X_{(n)} \leq 6$ 。又 $X_{(n)}$ 太小也應拒絕 H_0 , 如 $X_{(n)} < 4.6$, 最後即使 $X_{(n)}$ 略小於6, 也顯示 H_0 可能不真。不過門檻應與 n 有關。在本例裡, 我們取拒絕域為

$$(2.13) \quad C_1 = \{X_{(n)} < 4.6, \text{ 或 } X_{(n)} > 5.9\}。$$

則檢力函數為

$$K_1(\theta) = P_\theta(X_{(n)} < 4.6) + P_\theta(X_{(n)} > 5.9)。$$

其中

$$\begin{aligned} P_\theta(X_{(n)} < 4.6) &= \begin{cases} 1 & , \theta \leq 4.6, \\ (4.6/\theta)^n & , \theta > 4.6, \end{cases} \\ P_\theta(X_{(n)} > 5.9) &= 1 - P_\theta(X_{(n)} \leq 5.9) \\ &= \begin{cases} 0 & , \theta \leq 5.9, \\ 1 - (5.9/\theta)^n & , \theta > 5.9。 \end{cases} \end{aligned}$$

因此

$$(2.14) \quad K_1(\theta) = \begin{cases} 1 & , \theta \leq 4.6, \\ (4.6/\theta)^n & , 4.6 < \theta \leq 5.9, \\ (4.6/\theta)^n + 1 - (5.9/\theta)^n & , \theta > 5.9。 \end{cases}$$

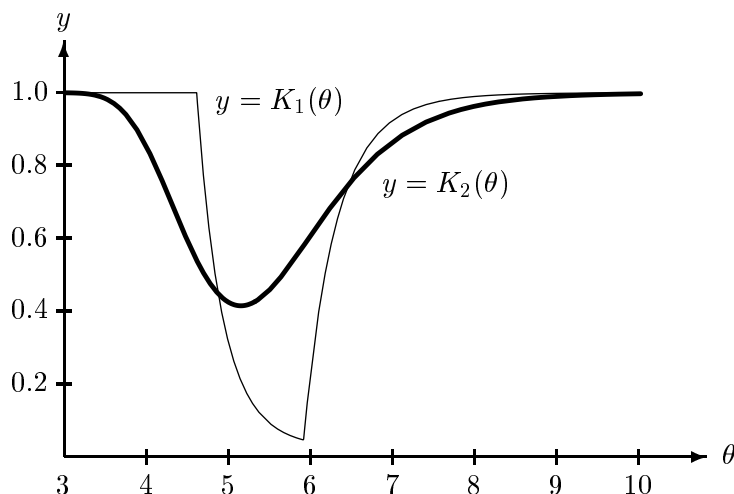
圖2.6 例2.11中對應拒絕域 C_1 及 C_2 之檢力函數

圖2.6給出 $n = 15$ 時, $y = K_1(\theta)$ 之圖形。其圖形在 $4.6 < \theta \leq 5.9$ 時為漸減, 在 $\theta > 5.9$ 時, 為漸增。

其次考慮基於樣本平均 $\bar{X}_n$ 之檢定。由於 $E(\bar{X}_n) = \theta/2$, 欲與前述檢定比較, 取拒絕域

$$(2.15) \quad C_2 = \{\bar{X}_n < 2.3, \text{ 或 } \bar{X}_n > 2.95\}。$$

$\bar{X}_n$ 的分佈雖可求出, 但較繁瑣。因 $\text{Var}(X) = \theta^2/12$, 由中央極限定理, n 較大時, $\bar{X}_n$ 有近似的 $\mathcal{N}(\theta/2, \theta^2/(12n))$ 分佈。因此得近似的檢力函數

$$\begin{aligned} (2.16) \quad K_2(\theta) &= P_\theta(\bar{X}_n < 2.3) + P_\theta(\bar{X}_n > 2.95) \\ &= 1 - P_\theta\left(\frac{2.3 - \theta/2}{\theta/\sqrt{12n}} \leq Z \leq \frac{2.95 - \theta/2}{\theta/\sqrt{12n}}\right) \\ &= 1 - \Phi\left(\frac{\sqrt{12n}(2.95 - \theta/2)}{\theta}\right) + \Phi\left(\frac{\sqrt{12n}(2.3 - \theta/2)}{\theta}\right)。 \end{aligned}$$

圖2.6中亦給出當 $n = 15$ 時, $y = K_2(\theta)$ 之圖形。

我們知道所謂“好的”檢定, 乃指在 H_0 之下, 檢力函數要愈小愈好, 而在 H_a 之下, 檢力函數要愈大愈好。由圖2.6可看出, 基於充分統計量 $X_{(n)}$ 之檢定, 比基於樣本平均 $\bar{X}_n$ 之檢定應是較好很多:

在 H_0 之下(即 $5 \leq \theta \leq 6$), 皆有 $K_1(\theta) < K_2(\theta)$ 。而且

$$\begin{aligned}
 \sup_{5 \leq \theta \leq 6} K_1(\theta) &= \max\{K_1(5), K_1(6)\} \\
 &= \max\left\{\left(\frac{4.6}{5}\right)^{15}, \left(\frac{4.6}{6}\right)^{15} + 1 - \left(\frac{5.9}{6}\right)^{15}\right\} \\
 &\doteq \max\{0.2863, 0.2414\} = 0.2863, \\
 \sup_{5 \leq \theta \leq 6} K_2(\theta) &= \max\{K_2(5), K_2(6)\} \\
 &\doteq \max\{1 - \Phi(1.207) + \Phi(-0.537), 1 - \Phi(-0.112) + \Phi(-1.566)\} \\
 &\doteq \max\{1 - 0.8863 + 0.2957, 1 - 0.4554 + 0.0587\} \\
 &= \max\{0.4094, 0.6033\} = 0.6033。
 \end{aligned}$$

至於對 $\theta < 5$ 或 $\theta > 6$, 由圖2.6可看出, 除了在二不太大的區間(5的左側及6的右側), 皆有 $K_1(\theta) > K_2(\theta)$ 。

最後, 一檢定之檢力函數若滿足

$$(2.17) \quad K(\theta_1) \geq K(\theta_2), \forall \theta_1 \in \Omega_0^c, \theta_2 \in \Omega_0,$$

便稱此檢定為**不偏的**。不偏是“合理的”檢定所該具備的條件。例2.6是檢定 $H_0: \theta \leq 0.5$, vs. $H_a: \theta > 0.5$, 而 $K_1(\theta)$ 及 $K_2(\theta)$ 皆為 θ 之增函數, 因此二檢定皆為不偏的。例2.7是檢定 $H_0: \mu \leq \mu_0$, vs. $H_a: \mu > \mu_0$, 而 $K(\mu)$ 為 μ 之漸增函數, 因此亦為不偏的。由圖2.4及2.5也可看出此點。至於在例2.11中, 當 $n = 15$ 時, 二檢定皆非不偏的。不過可適當的調整拒絕域, 使檢定成為不偏的(習題第14題)。

8.3 假設檢定之進一步探討

上一節我們討論的都是檢定單一母體的期望值 μ , 且除了例2.6、2.7及2.11, 虛無假設都是簡單假設。當對立假設為 $H_a: \mu > \mu_0$ ($\mu < \mu_0$), 虛無假設為 $H_0: \mu \leq \mu_0$ ($\mu \geq \mu_0$), 與 $H_0: \mu = \mu_0$ 時, 會有相同的拒絕域。這要稍微想一下才會明白。本節我們來看一些其他的檢定。

8.3.1 期望值之差的檢定

我們先來看二期望值之差的檢定。

在很多實際的應用中，常會遇到兩母體期望值之差的檢定問題。譬如說，我們可能會想知道上過某種特別的教學法，學生之平均成績，是否會較未上過該種教學法的學生平均成績高？或是否高出至少3分？

假設有兩組獨立的隨機樣本 $X_1, \dots, X_{n_1}$ ，及 $Y_1, \dots, Y_{n_2}$ ，分別有 $\mathcal{N}(\mu_1, \sigma_1^2)$ 及 $\mathcal{N}(\mu_2, \sigma_2^2)$ 分佈，樣本數各為 n_1, n_2 。我們要檢定 $H_0: \mu_2 - \mu_1 = \delta$ ，其中 δ 為一給定之常數，而對立假設可能為 $H_a: \mu_2 - \mu_1 \neq \delta, \mu_2 - \mu_1 > \delta$ ，或 $\mu_2 - \mu_1 < \delta$ 三種之一。底下為一些預備知識，其證明我們留在習題中(第1、2題)：令 $\bar{X}, S_1^2$ 及 $\bar{Y}, S_2^2$ 分別表第一組及第二組樣本之樣本平均及樣本變異數。則

$$(3.1) \quad Z = \frac{(\bar{Y} - \bar{X}) - (\mu_2 - \mu_1)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

有 $\mathcal{N}(0, 1)$ 分佈。又

$$V = \frac{(n_1 - 1)S_1^2}{\sigma_1^2} + \frac{(n_2 - 1)S_2^2}{\sigma_2^2} \sim \chi_{n_1+n_2-2}^2。$$

仿第四章例3.2，得

$$T_1 = \frac{Z}{\sqrt{V/(n_1 + n_2 - 2)}} \sim \mathcal{T}_{n_1+n_2-2}。$$

若 σ_1^2, σ_2^2 未知且不相等，則 T 並非統計量。但若 $\sigma_1^2 = \sigma_2^2$ ，則 $T_1 = T$ ，其中

$$(3.2) \quad T = \frac{(\bar{Y} - \bar{X}) - (\mu_2 - \mu_1)}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

為一有 $\mathcal{T}_{n_1+n_2-2}$ 分佈之統計量。

現設 σ_1^2, σ_2^2 為已知。則對 $H_a: \mu_2 - \mu_1 \neq \delta, \mu_2 - \mu_1 > \delta, \mu_2 - \mu_1 < \delta$ ，在型 I 錯誤機率為 α 之下，拒絕域分別為 $\{|Z_0| > z_{1-\alpha/2}\}$ ， $\{Z_0 > z_{1-\alpha}\}$ ，及 $\{Z_0 < -z_{1-\alpha}\}$ ，其中

$$(3.3) \quad Z_0 = \frac{\bar{Y} - \bar{X} - \delta}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}。$$

例3.1 假設兩種香菸之尼古丁(nicotine)含量(單位為毫克)分別有 $\mathcal{N}(\mu_1, 1.2^2)$ 及 $\mathcal{N}(\mu_2, 1.4^2)$ 分佈, 欲檢驗 $H_0: \mu_2 - \mu_1 = 2$, vs. $H_a: \mu_2 - \mu_1 \neq 2$ 。第一種香菸取樣50, 得 $\bar{X} = 23.8$, 第二種香菸取樣40, 得 $\bar{Y} = 26.1$ 。則有

$$Z_0 = \frac{26.1 - 23.8 - 2}{\sqrt{1.2^2/50 + 1.4^2/40}} \doteq 1.076。$$

因 $|Z_0| < z_{0.975} = 1.96$, 故在 $\alpha = 0.05$ 之下, 不能拒絕 H_0 。事實上, 由於 p -值 $\doteq 0.2820$, 所以只要 $\alpha < 0.2820$, 皆不能拒絕 H_0 。

若兩組獨立的隨機樣本, 其分佈不一定皆是常態, 則只要樣本數 n_1, n_2 皆夠大, 利用中央極限定理, 先得(3.1)式中的 Z 有近似的 $\mathcal{N}(0, 1)$ 分佈, 便可仿例2.3的作法。即使變異數不知道, 仍可仿例2.5, 以 S_1^2 及 S_2^2 分別取代 σ_1^2 及 σ_2^2 。

其次我們來看 σ_1^2, σ_2^2 未知但相等的情況, 即設 $\sigma_1^2 = \sigma_2^2$ 。則在型 I 錯誤機率為 α 之下, 檢定 $H_0: \mu_2 - \mu_1 = \delta$, vs. $H_a: \mu_2 - \mu_1 \neq \delta$, $\mu_2 - \mu_1 > \delta$, 或 $\mu_2 - \mu_1 < \delta$, 拒絕域分別為 $\{|T_0| > t_{1-\alpha/2, n_1+n_2-2}\}$, $\{T_0 > t_{1-\alpha, n_1+n_2-2}\}$, 及 $\{T_0 < -t_{1-\alpha, n_1+n_2-2}\}$, 其中

$$(3.4) \quad T_0 = \frac{\bar{Y} - \bar{X} - \delta}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim \mathcal{T}_{n_1+n_2-2}。$$

這種檢定稱為**二樣本之 t 檢定**(two-sample t -test)。

例3.2 假設欲比較兩組學生之某種表現。取12位第一組的學生, 測驗後得平均成績81, 樣本標準差4; 取10位第二組的學生, 測驗後得平均成績85, 樣本標準差5。又設兩組學生之成績皆有常態分佈, 且變異數相同。欲檢驗 $H_0: \mu_2 - \mu_1 = 2$, vs. $H_a: \mu_2 - \mu_1 > 2$ 。

首先由所給數據得

$$T_0 = \frac{85 - 81 - 2}{\sqrt{\frac{11 \cdot 16 + 9 \cdot 25}{12 + 10 - 2}} \sqrt{\frac{1}{12} + \frac{1}{10}}} \doteq 1.043。$$

我們要查 $\mathcal{T}_{20}$ 分佈之機率值。由附表6, 利用內插法得 p -值約為(不過因表中所給的機率值兩兩間不是很接近(如0.80, 0.85, 相差0.05), 所以誤差可能

不會太小)

$$p \doteq P(T_0 > 1.043) \doteq 0.155。$$

故只要型 I 錯誤之機率 α 小於 0.155 (這是大部分的情況), 皆不能拒絕 $\mu_2 - \mu_1 = 2$ 之假設。

讀者可自行練習, 若改為檢定 $H_0: \mu_2 - \mu_1 = 0.5$, vs. $H_a: \mu_2 - \mu_1 > 0.5$, 其餘條件不變, 則得 $T_0 \doteq 1.826$, 且 p -值約為 $P(T_0 > 1.826) \doteq 0.043$ 。故若 $\alpha = 0.05$, 則得不能接受 $H_0: \mu_2 - \mu_1 = 0.5$ 之假設。

若 σ_1^2, σ_2^2 不知, 且 $\sigma_1^2 \neq \sigma_2^2$ (或不知 σ_1^2 與 σ_2^2 是否相等), 我們介紹一簡單的做法。先取 n_1, n_2 中較小者, 假設為 n_1 。將兩組樣本隨機地配成 n_1 對, 並考慮其差。此 n_1 個差, 形成一組 i.i.d. 的隨機變數, 以 $\mathcal{N}(\mu_2 - \mu_1, \sigma_1^2 + \sigma_2^2)$ 為共同分佈。在 $H_0: \mu_2 - \mu_1 = \delta$ 之下, 其共同分佈為期望值 $\mu = \delta$ 之常態分佈。因此可以上一節單樣本 t 檢定的方法檢定 H_0 (見例 2.4)。

至於若兩組隨機樣本並不獨立, 就要修正檢定的方法, 不過在此不擬討論。

8.3.2 變異數的檢定

母體變異數之檢定也是很重要的, 期望值如果是目標值, 則變異數量測數據偏離目標值之情況。所以不論是有關產品之品質, 或考題的製作, 都會很在乎變異數的大小。對於產品的品質, 通常會希望變異數不要超過某一定值。至於對一特定的測驗, 為了有鑑別效果, 有時會預設變異數, 或希望變異數超過某一定值。另外, 兩種不同的方法所產生的變異數也是常需要比較的。

底下我們只考慮母體有常態分佈。設 $X_1, \dots, X_n$ 為一組隨機樣本, 以 $\mathcal{N}(\mu, \sigma^2)$ 為其共同分佈。欲檢定 $H_0: \sigma^2 = \sigma_0^2$ 。如以往, 對立假設可為 $H_a: \sigma^2 \neq \sigma_0^2, \sigma^2 > \sigma_0^2$, 或 $\sigma^2 < \sigma_0^2$ 。首先設 μ 為已知。則當 H_0 為真時,

$$L = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma_0^2} \sim \chi_n^2,$$

其中 n 為樣本數。在型 I 錯誤機率為 α 之下, 拒絕域可分別取成 $\{L > \chi_{1-\alpha/2, n}^2 \text{ 或 } L < \chi_{\alpha/2, n}^2\}$, $\{L > \chi_{1-\alpha, n}^2\}$, 及 $\{L < \chi_{\alpha, n}^2\}$ 。此處 $\chi_{p, m}^2$ 為 χ_m^2 之第 p 分位數。即若 U 有 χ_m^2 分佈, $m \geq 1$, 則對 $\forall 0 < p < 1$, $\chi_{p, m}^2$ 滿足

$$(3.5) \quad P(U \leq \chi_{p, m}^2) = p.$$

其次設 μ 為未知。仍令 S^2 表樣本變異數。由第四章定理3.1, 當 H_0 為真時,

$$K = \frac{(n-1)S^2}{\sigma_0^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma_0^2} \sim \chi_{n-1}^2.$$

在型 I 錯誤機率為 α 之下, 拒絕域可分別取成 $\{K > \chi_{1-\alpha/2, n-1}^2 \text{ 或 } K < \chi_{\alpha/2, n-1}^2\}$, $\{K > \chi_{1-\alpha, n-1}^2\}$, 及 $\{K < \chi_{\alpha, n-1}^2\}$ 。

一般如果沒特別說明, 就表 μ 為未知。

例3.3 某製造商宣稱其生產的汽車電池之壽命有常態分佈, 標準差 $S = 0.9$ (年)。隨機地取10個電池測試, 得樣本標準差 $S = 1.2$ (年)。據此是否可判定電池壽命之標準差 $\sigma > 0.9$? 試以 $\alpha = 0.05$ 做一檢定。

解.依題意, 設 μ 為未知。在 $\alpha = 0.05$ 之下, 欲檢定 $H_0: \sigma^2 = 0.9^2$, vs. $H_a: \sigma^2 > 0.9^2$ 。由附表5, $\chi_{0.9, 9}^2 = 16.92$, 而

$$K = \frac{(n-1)S^2}{\sigma_0^2} = \frac{9 \cdot 1.2^2}{0.9^2} = 16.0 < 16.92,$$

故在 $\alpha = 0.05$ 之下, 無法拒絕 $\sigma = 0.9$ 。

我們也可求出 p -值 $\doteq 0.07$ 。所以如果取 $\alpha = 0.1$, 則會拒絕 H_0 。

最後, 在同樣得到 $S = 1.2$ 之下, 究竟樣本數較大或較小, 會使 H_0 較易被拒絕? 此點留給讀者自行思考。

設有兩組獨立的隨機樣本 $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$, 分別有 $\mathcal{N}(\mu_1, \sigma_1^2)$ 及 $\mathcal{N}(\mu_2, \sigma_2^2)$ 分佈。又以 S_1^2, S_2^2 分別表兩組樣本之樣本變異數。欲檢定 $H_0: \sigma_1^2 = \sigma_2^2$ 。對立假設為 $H_a: \sigma_1^2 \neq \sigma_2^2$, $H_a: \sigma_1^2 > \sigma_2^2$, 或 $H_a: \sigma_1^2 < \sigma_2^2$ 。

首先設 μ_1, μ_2 皆為已知。因 $\sum_{i=1}^{n_1}(X_i - \mu_1)^2/\sigma_1^2$ 及 $\sum_{j=1}^{n_2}(Y_j - \mu_2)^2/\sigma_2^2$ 分別有 $\chi_{n_1}^2$ 及 $\chi_{n_2}^2$ 分佈，依第四章定義3.2，

$$\frac{\sum_{i=1}^{n_1}(X_i - \mu_1)^2/(n_1\sigma_1^2)}{\sum_{j=1}^{n_2}(Y_j - \mu_2)^2/(n_2\sigma_2^2)} \sim \mathcal{F}_{n_1, n_2},$$

而 H_0 為真時，

$$F = \frac{\sum_{i=1}^{n_1}(X_i - \mu_1)^2/n_1}{\sum_{j=1}^{n_2}(Y_j - \mu_2)^2/n_2} \sim \mathcal{F}_{n_1, n_2} \circ$$

在型 I 錯誤機率為 α 之下，對三個不同的 H_a ，拒絕域可分別取成

$$\begin{aligned} & \{F > f_{1-\alpha/2, n_1, n_2}, \text{ 或 } F < f_{\alpha/2, n_1, n_2}\} \\ & = \{F > f_{1-\alpha/2, n_1, n_2}, \text{ 或 } 1/F > f_{1-\alpha/2, n_2, n_1}\} \end{aligned}$$

(等號右側是將拒絕域皆表為 $\mathcal{F}$ 分佈之右端機率，因附表7只給出右端機率)，

$$\{F > f_{1-\alpha, n_1, n_2}\}, \text{ 及 } \{1/F > f_{1-\alpha, n_2, n_1}\} \circ$$

此處 $f_{p, n, m}$ 為 $\mathcal{F}_{n, m}$ 分佈之第 p 分位數。即若隨機變數 V 有 $\mathcal{F}_{n, m}$ 分佈， $n, m \geq 1$ ，則對 $\forall 0 < p < 1$ ， $f_{p, n, m}$ 滿足

$$(3.6) \quad P(V \leq f_{p, n, m}) = p \circ$$

其次設 μ_1, μ_2 皆為未知。由第四章例3.4，

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim \mathcal{F}_{n_1-1, n_2-1} \circ$$

而 H_0 為真時，

$$\frac{S_1^2}{S_2^2} \sim \mathcal{F}_{n_1-1, n_2-1} \circ$$

由此對三個不同的 H_a ，可分別得到對應的拒絕域：

$$\begin{aligned} & \{S_1^2/S_2^2 > f_{1-\alpha/2, n_1-1, n_2-1}, \text{ 或 } S_1^2/S_2^2 < f_{\alpha/2, n_1-1, n_2-1}\}, \\ & \{S_1^2/S_2^2 > f_{1-\alpha, n_1-1, n_2-1}\}, \text{ 及 } \{S_2^2/S_1^2 > f_{1-\alpha, n_2-1, n_1-1}\} \circ \end{aligned}$$

註3.3 當 V 有 $\mathcal{F}_{n,m}$ 分佈時, $1/V$ 有 $\mathcal{F}_{m,n}$ 分佈(4.3節習題第7題)。因此

$$(3.7) \quad f_{p,n,m} = \frac{1}{f_{1-p,m,n}}.$$

附表7只給較大的 p 之 $f_{p,n,m}$, 對較小的 p , 就要利用上式以求出 $f_{p,n,m}$ 。

例3.4 欲比較兩種不同結構的鋼之張力強度。經試驗後得 $n_1 = 13$, $S_1^2 = 19.2$, $n_2 = 16$, $S_2^2 = 3.5$, 單位我們略去。假設張力皆有常態分佈, 變異數分別為 σ_1^2 及 σ_2^2 。試在 $\alpha = 0.02$ 之下, 檢定 $H_0: \sigma_1^2 = \sigma_2^2$, vs. $H_a: \sigma_1^2 \neq \sigma_2^2$ 。

解. 依題意, μ_1, μ_2 皆設為未知。拒絕域為 $\{S_1^2/S_2^2 > f_{0.99,12,15}, \text{ 或 } S_1^2/S_2^2 < f_{0.01,12,15}\}$ 。由附表7得 $f_{0.99,12,15} \doteq 3.67$, 而

$$f_{0.01,12,15} = \frac{1}{f_{0.99,15,12}} \doteq \frac{1}{4.01} \doteq 0.2493.$$

由於

$$\frac{S_1^2}{S_2^2} = \frac{19.2}{3.5} \doteq 5.49 > 3.67,$$

故在 $\alpha = 0.02$ 之下, 拒絕 $\sigma_1^2 = \sigma_2^2$ 。

8.3.3 二比例之檢定

二比例是否相等, 常是我們所關切的。例如, 想知道抽煙者得肺癌之比例, 是否較不抽煙者高? 這問題相當於有兩組獨立的伯努力分佈的母體, 樣本數分別為 n_1 及 n_2 , 每次成功之機率分別為 p_1 及 p_2 。欲檢定 $H_0: p_1 = p_2$, 對立假設可取為 $H_a: p_1 \neq p_2$, $p_1 > p_2$, 或 $p_1 < p_2$ 。當然這等價於 $H_0: p_1 - p_2 = 0$, vs. $H_a: p_1 - p_2 \neq 0$, $p_1 - p_2 > 0$, 與 $p_1 - p_2 < 0$ 。令 X_1, X_2 分別表兩組樣本之成功數, 則 X_1 有 $\mathcal{B}(n_1, p_1)$ 分佈, X_2 有 $\mathcal{B}(n_2, p_2)$ 分佈。

當 n_1, n_2 皆很大, p_1, p_2 的估計量 $\hat{p}_1 = X_1/n_1$, $\hat{p}_2 = X_2/n_2$, 分別有近似的 $\mathcal{N}(p_1, p_1(1-p_1)/n_1)$, 及 $\mathcal{N}(p_2, p_2(1-p_2)/n_2)$ 分佈, 而且 $\hat{p}_1$ 與 $\hat{p}_2$ 獨立。因此

$$B = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}}$$

有近似的 $\mathcal{N}(0, 1)$ 分佈。當 H_0 成立時, 即 $p_1 = p_2 = p$, B 成為

$$B_1 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{p(1-p)(1/n_1 + 1/n_2)}}。$$

但 B_1 並非一統計量，因其中含有未知的量 p 。將兩組樣本併在一起，得 p 之估計量

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}。$$

將 B_1 中的 p 以 $\hat{p}$ 代入，得

$$B_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}}。$$

B_0 有近似的 $\mathcal{N}(0, 1)$ 分佈。因此在型 I 錯誤機率為 α 之下，拒絕域可分別取為 $\{|B_0| > z_{1-\alpha/2}\}$ ， $\{B_0 > z_{1-\alpha}\}$ 及 $\{B_0 < -z_{1-\alpha}\}$ 。

例3.5 欲調查高雄縣民及高雄市民對高雄縣市之合一的支持度是否相同。訪問了200位高雄縣民，及500位高雄市民。前者有120位支持，後者有240位支持。試在 $\alpha = 0.025$ 之下，判定高雄縣民對縣市合一之支持度是否較高。

解.令 p_1 及 p_2 分別表高雄縣民及高雄市民之支持比例。欲檢定 $H_0 : p_1 = p_2$, vs. $H_a : p_1 > p_2$, $\alpha = 0.025$ 。因

$$\hat{p}_1 = \frac{120}{200} = 0.6, \quad \hat{p}_2 = \frac{240}{500} = 0.48, \quad \hat{p} = \frac{120 + 240}{200 + 500} \doteq 0.514,$$

故

$$B_0 \doteq \frac{0.6 - 0.48}{\sqrt{0.514 \cdot 0.486 \cdot (1/200 + 1/500)}} \doteq 2.87。$$

又 p -值 $= P(Z > 2.87) \doteq 0.0021 < 0.025$ 。因此拒絕 H_0 ，而接受 $p_1 > p_2$ 。即接受高雄縣民支持縣市合一的比例高於高雄市民。

8.4 概似比檢定

在上兩節我們介紹了一些簡單的檢定程序。如同在點估計裡，檢定中亦有很多不同的方法。本節我們介紹一常見的方法，即**概似比檢**

定(likelihood ratio test, 簡稱LRT)。此法與第六章的**最大概似法**(MLE)關係密切, 並且都是通常可行且不錯的方法, 只是後者針對點估計, 前者適用於假設檢定。

設 $X_1, \dots, X_n$ 為一組由p.d.f. $f(x|\theta)$ 所產生之隨機樣本, 其中 θ 可能為一向量。則如同第六章所定義的, 概似函數為

$$(4.1) \quad L(\theta|\mathbf{x}) = L(\theta|x_1, \dots, x_n) = \prod_{i=1}^n f(x_i|\theta) = f(\mathbf{x}|\theta), \theta \in \Omega,$$

其中 Ω 表整個參數空間。欲檢定 $H_0: \theta \in \Omega_0$, vs. $H_a: \theta \in \Omega_0^c$, 概似比檢定之統計量為

$$(4.2) \quad \lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Omega_0} L(\theta|\mathbf{x})}{\sup_{\theta \in \Omega} L(\theta|\mathbf{x})}。$$

一檢定若拒絕域有 $\{\lambda(\mathbf{X}) \leq c\}$ 之形式, 其中 $0 \leq c \leq 1$, 便稱為**概似比檢定**。必須要一提的是, 一般統計量為一隨機變數, 但在(4.2)式中, 我們是用小寫的 $\mathbf{x}$, 符號在此處有點混淆。這是因在 $L(\theta|\mathbf{x})$ 中, $\mathbf{x}$ 是以小寫表示。底下也仍有類似的情況。不過在涉及機率時(如底下(4.3)式), 就須用大寫的 $\mathbf{X}$ 。

我們以離散型的p.d.f.來解釋LRT。設 $f(x|\theta)$ 為一離散型的p.d.f.。則 $\lambda(\mathbf{x})$ 之分子, 為參數 θ 在虛無假設下(即 $\theta \in \Omega_0$), 會使 $\mathbf{x}$ 被觀測到之最大機率; $\lambda(\mathbf{x})$ 之分母, 則為參數 θ 在整個參數空間中(即 $\theta \in \Omega$), 會使 $\mathbf{x}$ 被觀測到之最大機率。此二最大機率之比值若較小, 表在對立假設中, 存在某一參數, 比任一虛無假設中的參數, 都使觀測值 $\mathbf{x}$ 更易產生。在這情況下, LRT便判定 H_0 該拒絕, 而 H_a 該接受。至於常數 c 的選擇, 則由顯著水準決定。即對任一給定的 $0 \leq \alpha \leq 1$, c 滿足

$$(4.3) \quad \sup_{\theta \in \Omega_0} P_{\theta}(\lambda(\mathbf{X}) \leq c) = \alpha。$$

可看出LRT的想法與MLE近似。並且若以 $\hat{\theta} = \hat{\theta}(\mathbf{x})$ 表 θ 之MLE, 即

$$L(\hat{\theta}|\mathbf{x}) = \max\{L(\theta|\mathbf{x}), \theta \in \Omega\},$$

且以 $\hat{\theta}_0 = \hat{\theta}_0(\mathbf{x})$ 表 θ 在 Ω_0 中之MLE, 即

$$L(\hat{\theta}_0|\mathbf{x}) = \max\{L(\theta|\mathbf{x}), \theta \in \Omega_0\},$$

則LRT統計量成爲

$$\lambda(\mathbf{x}) = \frac{L(\hat{\theta}_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})}。$$

在絕大部分的情況中, $0 < c < 1$, c 不會取0或1。此因由(4.3)式可看出, $c = 0$ 對應顯著水準 $\alpha = 0$; $c = 1$ 則對應 $\alpha = 1$, 而通常 α 不會取爲0或1。

例4.1 設 $X_1, \dots, X_n$ 爲一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本, 其中 σ^2 爲已知。欲檢定 $H_0: \mu = \mu_0$, vs. $H_a: \mu \neq \mu_0$ 。由於 $\Omega_0 = \{\mu_0\}$ 中只有一元素, LRT統計量 $\lambda(\mathbf{x})$ 之分子爲 $L(\mu_0|\mathbf{x})$ 。又由第六章例3.6, μ 之MLE爲 $\bar{X}$ (對樣本平均 $\bar{X}_n$, 我們省略足標 n , 亦以 $\bar{x}$ 取代 $\bar{x}_n$)。因此LRT統計量

$$\begin{aligned} \lambda(\mathbf{x}) &= \frac{(2\pi)^{-n/2} \sigma^{-n} \exp\{-\sum_{i=1}^n (x_i - \mu_0)^2 / (2\sigma^2)\}}{(2\pi)^{-n/2} \sigma^{-n} \exp\{-\sum_{i=1}^n (x_i - \bar{x})^2 / (2\sigma^2)\}} \\ &= \exp\left\{-\sum_{i=1}^n (x_i - \mu_0)^2 + \sum_{i=1}^n (x_i - \bar{x})^2\right\} / (2\sigma^2) \\ &= \exp\{-n(\bar{x} - \mu_0)^2 / (2\sigma^2)\}。 \end{aligned}$$

此處用到(見第四章定理1.1)

$$\sum_{i=1}^n (x_i - \mu_0)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2。$$

故LRT檢定之拒絕域爲

$$\{\lambda(\mathbf{X}) \leq c\} \Leftrightarrow \{|\bar{X} - \mu_0| \geq c_1\},$$

其中 $0 < c < 1$, 且

$$c_1 = (-(2\sigma^2 \log c)/n)^{1/2},$$

$c_1 \in (0, \infty)$ 。給定一顯著水準 α , 因 Ω_0 中只有一個 μ_0 , 常數 c_1 由

$$\sup_{\mu \in \Omega_0} P_\mu(|\bar{X} - \mu_0| \geq c_1) = P_{\mu_0}(|\bar{X} - \mu_0| \geq c_1) = \alpha$$

決定。又當 $\mu = \mu_0$, $\sqrt{n}(\bar{X} - \mu_0)/\sigma$ 有 $\mathcal{N}(0, 1)$ 分佈, 故 c_1 可取成 $z_{1-\alpha/2}\sigma/\sqrt{n}$ 。此與 8.2 節中, 以直觀的方法所得之拒絕域相同(見例 2.3 之後的討論)。

例 4.2 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x|\theta) = e^{-(x-\theta)}$, $x \geq \theta$, 所產生之隨機樣本, 其中 $\theta \in R$ 。欲檢定 $H_0 : \theta \leq \theta_0$, vs. $H_a : \theta > \theta_0$, 其中 θ_0 為某一定值。

令 $\Omega_0 = \{\theta | \theta \leq \theta_0\}$ 。概似函數為

$$L(\theta|\mathbf{x}) = \exp\left\{-\sum_{i=1}^n x_i + n\theta\right\}, \theta \leq x_{(1)}。$$

由於 $L(\theta|\mathbf{x})$ 為 θ 之漸增函數, $-\infty < \theta \leq x_{(1)}$, 所以 $\lambda(\mathbf{x})$ 之分母為 $L(x_{(1)}|\mathbf{x}) = \exp\{-\sum_{i=1}^n x_i + nx_{(1)}\}$ 。若 $x_{(1)} \leq \theta_0$, 則對 $\theta \in \Omega_0$, $\theta \leq x_{(1)} \leq \theta_0$, 因此 $\lambda(\mathbf{x})$ 之分子亦為 $L(x_{(1)}|\mathbf{x})$ 。若 $x_{(1)} > \theta_0$, 則對 $\theta \in \Omega_0$, $\theta \leq \theta_0 < x_{(1)}$, 因此 $\lambda(\mathbf{x})$ 之分子為 $L(\theta_0|\mathbf{x}) = \exp\{-\sum_{i=1}^n x_i + n\theta_0\}$ 。故得

$$\lambda(\mathbf{x}) = \begin{cases} 1 & , \text{若 } x_{(1)} \leq \theta_0, \\ \exp\{-n(x_{(1)} - \theta_0)\} & , \text{若 } x_{(1)} > \theta_0。 \end{cases}$$

圖 4.1 給出 $\lambda(\mathbf{x})$ 之圖形。

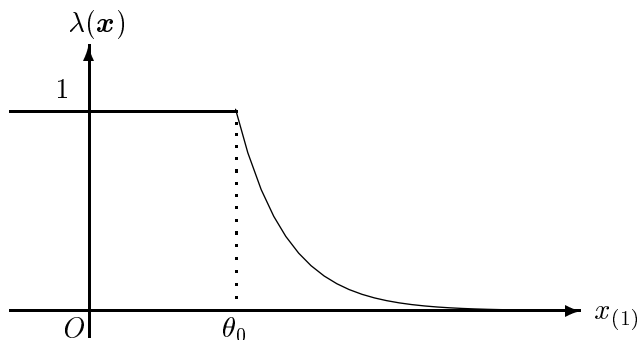


圖 4.1 例 4.2 中 $\lambda(\mathbf{x})$ 之圖形

LRT 檢定之拒絕域為

$$\{\lambda(\mathbf{X}) \leq c\} \Leftrightarrow \{X_{(1)} \geq c_1\},$$

其中 $0 < c < 1$, 且 $c_1 = \theta_0 - \log c/n$, $c_1 \in (\theta_0, \infty)$ 。給定一顯著水準 α , 常數 c_1 可由

$$\sup_{\theta \in \Omega_0} P_\theta(X_{(1)} \geq c_1) = \sup_{\theta \leq \theta_0} P_\theta(X_{(1)} \geq c_1) = \alpha$$

決定。又因 $X_1, \dots, X_n$ 為 i.i.d. 之隨機變數, 故

$$\begin{aligned} \sup_{\theta \leq \theta_0} P_\theta(X_{(1)} \geq c_1) &= \sup_{\theta \leq \theta_0} (P_\theta(X \geq c_1))^n \\ &= \sup_{\theta \leq \theta_0} \exp\{-n(c_1 - \theta)\} \\ &= \exp\{-n(c_1 - \theta_0)\}。 \end{aligned}$$

其中如前 X 代表 $X_1, \dots, X_n$ 中的某一個。解出

$$c_1 = \theta_0 - (\log \alpha)/n。$$

例4.1中之參數 σ^2 與檢定的參數 μ 不相干, 稱為**擾亂參數**(nuisance parameter)。若 σ^2 設為未知, 則可能會得到不同的檢定量。我們留在例4.5再討論。

在例4.1中, LRT之拒絕域只與充分統計量 $\bar{X}$ 有關, 而在例4.2中, LRT之拒絕域也只與充分統計量 $X_{(1)}$ 有關, 這其實並非偶然。

設 $T(\mathbf{X})$ 為 θ 之一充分統計量, 且以 $g(t|\theta)$ 為 p.d.f.。考慮基於 T (及其概似函數 $L^*(\theta|t) = g(t|\theta)$) 之 LRT。令 $\lambda^*(t)$ 表基於 T 之概似檢定量。直觀上, 由於 $T(\mathbf{X})$ 包含所有關於 T 之資訊, 基於 T 之檢定, 似應與基於完整的樣本 $\mathbf{X}$ 一樣好。底下的定理便證實此猜測。

定理4.1 設 $T = T(\mathbf{X})$ 為 θ 之一充分統計量, $\lambda^*(t)$ 及 $\lambda(\mathbf{x})$ 分別表基於 T 及 $\mathbf{X}$ 之 LRT。則對每一樣本點 $\mathbf{x}$,

$$(4.4) \quad \lambda^*(T(\mathbf{x})) = \lambda(\mathbf{x})。$$

證明. 首先由分解定理(見第五章定理2.2),

$$f(\mathbf{x}|\theta) = g(T(\mathbf{x})|\theta)h(\mathbf{x}),$$

其中 $g(t|\theta)$ 取為 T 之p.d.f., 而 $h(\mathbf{x})$ 與 θ 無關。因此

$$\begin{aligned}
 \lambda(\mathbf{x}) &= \frac{\sup_{\theta \in \Omega_0} L(\theta|\mathbf{x})}{\sup_{\theta \in \Omega} L(\theta|\mathbf{x})} \\
 &= \frac{\sup_{\theta \in \Omega_0} f(\mathbf{x}|\theta)}{\sup_{\theta \in \Omega} f(\mathbf{x}|\theta)} \\
 &= \frac{\sup_{\theta \in \Omega_0} g(T(\mathbf{x})|\theta)h(\mathbf{x})}{\sup_{\theta \in \Omega} g(T(\mathbf{x})|\theta)h(\mathbf{x})} \\
 &= \frac{\sup_{\theta \in \Omega_0} g(T(\mathbf{x})|\theta)}{\sup_{\theta \in \Omega} g(T(\mathbf{x})|\theta)} \quad (h \text{ 與 } \theta \text{ 無關}) \\
 &= \frac{\sup_{\theta \in \Omega_0} L^*(\theta|T(\mathbf{x}))}{\sup_{\theta \in \Omega} L^*(\theta|T(\mathbf{x}))} \quad (g \text{ 爲 } T \text{ 之 p.d.f.}) \\
 &= \lambda^*(T(\mathbf{x}))。
 \end{aligned}$$

證畢。

在例4.1以及4.2中, 我們皆將所得的LRT統計量 $\lambda(\mathbf{x})$, 化成較簡單的形式, 事實上就是化成充分統計量的函數。定理4.1告訴我們其實可一開始就考慮充分統計量之概似函數, 所得LRT統計量將相同, 見下例。

例4.3 在例4.1中, μ 之充分統計量 $\bar{X}$ 有 $\mathcal{N}(\mu, \sigma^2/n)$ 分佈, 欲檢定 $H_0: \mu = \mu_0$, vs. $H_a: \mu \neq \mu_0$ 。基於 $\bar{X}$ 之LRT統計量為

$$\begin{aligned}
 \lambda(\bar{x}) &= \frac{(2\pi)^{-1/2}(\sigma/\sqrt{n})^{-1} \exp\{-(\bar{x} - \mu_0)^2/(2\sigma^2/n)\}}{(2\pi)^{-1/2}(\sigma/\sqrt{n})^{-1} \exp\{-(\bar{x} - \bar{x})^2/(2\sigma^2/n)\}} \\
 &= \exp\{-(\bar{x} - \mu_0)^2/(2\sigma^2/n)\},
 \end{aligned}$$

與例4.1中之 $\lambda(\mathbf{x})$ 完全一樣。

例4.4 在例4.2中, $X_{(1)}$ 為 θ 之充分統計量。 $X_{(1)}$ 之概似函數(即其p.d.f.)為

$$L^*(\theta|x_{(1)}) = ne^{-n(x_{(1)}-\theta)}, \theta \leq x_{(1)}。$$

繼續作下去, 得到的拒絕域, 將與例4.2中所得之拒絕域相同。

例4.5 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本, σ^2 設為未

知, $\theta = (\mu, \sigma^2)$ 。欲檢定 $H_0: \mu = \mu_0$, vs. $H_a: \mu \neq \mu_0$ 。則LRT統計量為

$$\lambda(\mathbf{x}) = \frac{\sup_{\{\mu, \sigma^2 | \mu = \mu_0, \sigma^2 > 0\}} L(\mu, \sigma^2 | \mathbf{x})}{\sup_{\{\mu, \sigma^2 | \mu \in R, \sigma^2 > 0\}} L(\mu, \sigma^2 | \mathbf{x})} = \frac{L(\mu_0, \hat{\sigma}_0^2 | \mathbf{x})}{L(\hat{\mu}, \hat{\sigma}^2 | \mathbf{x})},$$

其中 $\hat{\mu} = \bar{x}$, $\hat{\sigma}^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / n$ 分別為 μ 及 σ^2 之MLE(見第六章例3.6),

而 $\hat{\sigma}_0^2 = \sum_{i=1}^n (x_i - \mu_0)^2 / n$ 。故

$$\begin{aligned} \lambda(\mathbf{x}) &= \frac{(2\pi)^{-n/2} (\hat{\sigma}_0)^{-n} \exp\{-\sum_{i=1}^n (x_i - \mu_0)^2 / (2\hat{\sigma}_0^2)\}}{(2\pi)^{-n/2} (\hat{\sigma})^{-n} \exp\{-\sum_{i=1}^n (x_i - \bar{x})^2 / (2\hat{\sigma}^2)\}} \\ &= \frac{(\hat{\sigma}_0)^{-n} \exp\{-n/2\}}{(\hat{\sigma})^{-n} \exp\{-n/2\}} = \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right)^{-n/2}. \end{aligned}$$

因此拒絕域為

$$\{\lambda(\mathbf{X}) \leq c\} \Leftrightarrow \{\hat{\sigma}_0^2 / \hat{\sigma}^2 \geq c_1\},$$

其中 $c_1 = c^{-n/2}$ 。又

$$\hat{\sigma}_0^2 = \hat{\sigma}^2 + (\bar{X} - \mu_0)^2,$$

因此

$$\hat{\sigma}_0^2 / \hat{\sigma}^2 = 1 + (\bar{X} - \mu_0)^2 / \hat{\sigma}^2.$$

由於樣本變異數

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n}{n-1} \hat{\sigma}^2,$$

故

$$\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2} = 1 + \frac{n}{n-1} \frac{(\bar{X} - \mu_0)^2}{S^2} = 1 + \frac{1}{n-1} T_n^2,$$

其中

$$T_n = T_n(\mathbf{X}) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{S}.$$

即得 $\hat{\sigma}_0^2 / \hat{\sigma}^2$ 為 $|T_n|$ 之一單調漸增函數。故拒絕域可改寫為 $\{|T_n(\mathbf{X})| \geq c_2\}$ 。再度由第四章例3.2, 在 H_0 之下, T_n 有 $\mathcal{T}_{n-1}$ 分佈。當顯著水準為 α ,

$c_2 = t_{1-\alpha/2, n-1}$ 。此與8.2節(見例2.4之前的推導)討論單樣本之 t 檢定, 所得拒絕域相同。

例如, 設 $n = 25$, 且取顯著水準 $\alpha = 0.05$ 。則拒絕域為 $\{|T_n| \geq t_{0.975, 24} \doteq 2.064\}$ 。

例4.6 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\theta_1, \theta_3)$ 分佈所產生之隨機樣本, $Y_1, \dots, Y_m$ 為一組由 $\mathcal{N}(\theta_2, \theta_3)$ 分佈所產生之隨機樣本, 二組隨機樣本假設為相互獨立。又設 $\theta_1, \theta_2 \in R$, $\theta_3 > 0$ 皆為未知。即參數 $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$, 且 $\Omega = \{\boldsymbol{\theta} | -\infty < \theta_1, \theta_2 < \infty, \theta_3 > 0\}$ 。欲檢定 $H_0 : \theta_1 = \theta_2$, vs. $H_a : \theta_1 \neq \theta_2$ 。即 $\Omega_0 = \{\boldsymbol{\theta} | -\infty < \theta_1 = \theta_2 < \infty, \theta_3 > 0\}$ 。在 $n + m > 2$ 之假設下, 試求概似比檢定。

解. 首先概似函數為

$$L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y}) = (2\pi\theta_3)^{-(n+m)/2} \exp\left\{-\frac{\sum_{i=1}^n (x_i - \theta_1)^2 + \sum_{j=1}^m (y_j - \theta_2)^2}{2\theta_3}\right\}。$$

解

$$\frac{\partial \log L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})}{\partial \theta_1} = 0, \quad \frac{\partial \log L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})}{\partial \theta_2} = 0, \quad \frac{\partial \log L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})}{\partial \theta_3} = 0,$$

得 $\theta_1, \theta_2, \theta_3$ 之MLE為

$$\begin{aligned} \hat{\theta}_1 &= \frac{\sum_{i=1}^n x_i}{n} = \bar{x}, \\ \hat{\theta}_2 &= \frac{\sum_{j=1}^m y_j}{m} = \bar{y}, \\ \hat{\theta}_3 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{j=1}^m (y_j - \bar{y})^2}{n + m}, \end{aligned}$$

且

$$L(\hat{\boldsymbol{\theta}}|\mathbf{x}, \mathbf{y}) = \max\{L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y}), \boldsymbol{\theta} \in \Omega\} = (2\pi\hat{\theta}_3 e)^{-(n+m)/2}。$$

當 $\boldsymbol{\theta} \in \Omega_0$, 概似函數為

$$L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y}) = (2\pi\theta_3)^{-(n+m)/2} \exp\left\{-\frac{\sum_{i=1}^n (x_i - \theta_1)^2 + \sum_{j=1}^m (y_j - \theta_1)^2}{2\theta_3}\right\}。$$

解

$$\frac{\partial \log L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})}{\partial \theta_1} = 0, \quad \frac{\partial \log L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})}{\partial \theta_3} = 0, \boldsymbol{\theta} \in \Omega_0,$$

得在 Ω_0 中, θ_1, θ_3 之MLE爲

$$\begin{aligned} \hat{\theta}'_1 &= \frac{\sum_{i=1}^n x_i + \sum_{j=1}^m y_j}{n+m} = \frac{n\bar{x} + m\bar{y}}{n+m}, \\ \hat{\theta}'_3 &= \frac{\sum_{i=1}^n (x_i - \hat{\theta}'_1)^2 + \sum_{j=1}^m (y_j - \hat{\theta}'_1)^2}{n+m}, \end{aligned}$$

且

$$L(\hat{\boldsymbol{\theta}}_0|\mathbf{x}, \mathbf{y}) = (2\pi\hat{\theta}'_3 e)^{-(n+m)/2}。$$

因此LRT統計量爲

$$\lambda(\mathbf{x}, \mathbf{y}) = \frac{L(\hat{\boldsymbol{\theta}}_0|\mathbf{x}, \mathbf{y})}{L(\hat{\boldsymbol{\theta}}|\mathbf{x}, \mathbf{y})} = \left(\frac{\hat{\theta}_3}{\hat{\theta}'_3}\right)^{(n+m)/2}。$$

故LRT檢定之拒絕域爲

$$\{\lambda(\mathbf{X}, \mathbf{Y}) \leq c\} \Leftrightarrow \{\hat{\theta}_3/\hat{\theta}'_3 \leq c_1\},$$

其中 $0 < c < 1$, 且 $c_1 = c^{2/(n+m)}$ 。又

$$\begin{aligned} \frac{\hat{\theta}_3}{\hat{\theta}'_3} &= \left(\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{j=1}^m (Y_j - \bar{Y})^2 \right) \\ &\quad \cdot \left(\sum_{i=1}^n \left(X_i - \frac{n\bar{X} + m\bar{Y}}{n+m} \right)^2 + \sum_{j=1}^m \left(Y_j - \frac{n\bar{X} + m\bar{Y}}{n+m} \right)^2 \right)^{-1}, \end{aligned}$$

其中

$$\begin{aligned} \sum_{i=1}^n \left(X_i - \frac{n\bar{X} + m\bar{Y}}{n+m} \right)^2 &= \sum_{i=1}^n \left((X_i - \bar{X}) + \left(\bar{X} - \frac{n\bar{X} + m\bar{Y}}{n+m} \right) \right)^2 \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + n \left(\bar{X} - \frac{n\bar{X} + m\bar{Y}}{n+m} \right)^2 \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + \frac{nm^2(\bar{X} - \bar{Y})^2}{(n+m)^2}。 \end{aligned}$$

同理

$$\sum_{j=1}^m \left(Y_j - \frac{n\bar{X} + m\bar{Y}}{n+m} \right)^2 = \sum_{j=1}^m (Y_j - \bar{Y})^2 + \frac{n^2 m (\bar{X} - \bar{Y})^2}{(n+m)^2}。$$

故

$$\begin{aligned} \frac{\hat{\theta}_3}{\hat{\theta}_3'} &= \frac{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{j=1}^m (Y_j - \bar{Y})^2}{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{j=1}^m (Y_j - \bar{Y})^2 + (nm/(n+m))(\bar{X} - \bar{Y})^2} \\ &= \left[1 + \frac{(nm/(n+m))(\bar{X} - \bar{Y})^2}{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{j=1}^m (Y_j - \bar{Y})^2} \right]^{-1} \\ &= \left[1 + \frac{1}{n+m-2} T^2 \right]^{-1}, \end{aligned}$$

其中

$$T = \frac{(nm/(n+m))^{1/2}(\bar{X} - \bar{Y})}{((\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{j=1}^m (Y_j - \bar{Y})^2)/(n+m-2))^{1/2}}。$$

可證明在 H_0 之下, T 有 $\mathcal{T}_{n+m-2}$ 分佈(習題第2題)。故拒絕 H_0 , 若且唯若 $|T| \geq c_2$ 。而在顯著水準 α 之下, c_2 可取為 $t_{1-\alpha/2, n+m-2}$ 。例如, 設 $n = 10, m = 6, \alpha = 0.05$ 。則拒絕域為 $\{|T| \geq t_{0.95, 14} = 2.145\}$ 。

在上例中, 在檢定二獨立常態分佈之期望值是否相等時, 我們假設變異數未知但相等。下例則是檢定二變異數是否相等。

例4.7 在上例中, 假設 X 有 $\mathcal{N}(\theta_1, \theta_3)$ 分佈, Y 有 $\mathcal{N}(\theta_2, \theta_4)$ 分佈, $n > 1, m > 1$, 其餘不變。即 $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4)$, 且 $\Omega = \{\boldsymbol{\theta} | -\infty < \theta_1, \theta_2 < \infty, 0 < \theta_3, \theta_4 < \infty\}$ 。欲檢定 $H_0 : \theta_3 = \theta_4$, vs. $H_a : \theta_3 \neq \theta_4$ 。即 $\Omega_0 = \{\boldsymbol{\theta} | -\infty < \theta_1, \theta_2 < \infty, 0 < \theta_3 = \theta_4 < \infty\}$ 。仿上例, 可得

$$\begin{aligned} L(\hat{\boldsymbol{\theta}}_0 | \mathbf{x}, \mathbf{y}) &= (2\pi\hat{\theta}_3' e)^{-(n+m)/2}, \\ L(\hat{\boldsymbol{\theta}} | \mathbf{x}, \mathbf{y}) &= (2\pi e)^{-(n+m)/2} \hat{\theta}_3^{-n/2} \hat{\theta}_4^{-m/2}, \end{aligned}$$

其中

$$\hat{\theta}_3' = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{j=1}^m (y_j - \bar{y})^2}{n+m},$$

$$\hat{\theta}_3 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n},$$

$$\hat{\theta}_4 = \frac{\sum_{j=1}^m (y_j - \bar{y})^2}{m},$$

此處 $\hat{\theta}_3'$ 即上例中之 $\hat{\theta}_3$ 。故LRT統計量為

$$\begin{aligned}\lambda(\mathbf{x}, \mathbf{y}) &= \frac{L(\hat{\theta}_0 | \mathbf{x}, \mathbf{y})}{L(\hat{\theta} | \mathbf{x}, \mathbf{y})} = \frac{\hat{\theta}_3^{n/2} \hat{\theta}_4^{m/2}}{(\hat{\theta}_3')^{(n+m)/2}} \\ &= \frac{(\sum_{i=1}^n (x_i - \bar{x})^2 / n)^{n/2} (\sum_{j=1}^m (y_j - \bar{y})^2 / m)^{m/2}}{((\sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{j=1}^m (y_j - \bar{y})^2) / (n+m))^{(n+m)/2}} \\ &= \frac{(n+m)^{(n+m)/2}}{n^{n/2} m^{m/2}} \frac{(\sum_{i=1}^n (x_i - \bar{x})^2)^{n/2} (\sum_{j=1}^m (y_j - \bar{y})^2)^{m/2}}{(\sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{j=1}^m (y_j - \bar{y})^2)^{(n+m)/2}} \\ &= \frac{(n+m)^{(n+m)/2}}{n^{n/2} m^{m/2}} \left(\frac{1}{1 + (m-1)/(n-1)F^{-1}} \right)^{n/2} \\ &\quad \cdot \left(\frac{1}{1 + (n-1)/(m-1)F} \right)^{m/2},\end{aligned}$$

其中

$$F = \frac{\sum_{i=1}^n (X_i - \bar{X})^2 / (n-1)}{\sum_{j=1}^m (Y_j - \bar{Y})^2 / (m-1)}.$$

故LRT檢定之拒絕域為

$$\{\lambda(\mathbf{X}, \mathbf{Y}) \leq c\}.$$

在 H_0 之下, $\theta_3 = \theta_4$, 此時 F 有 $\mathcal{F}_{n-1, m-1}$ 分佈。由於 $\lambda(\mathbf{X}, \mathbf{Y})$ 並非 F 之單調函數, 由 $\lambda(\mathbf{X}, \mathbf{Y}) \leq c$, 一般而言, 不易解出一以 F 表示之區域(除非 $n = m$, 或 n, m 較小時)。不過在顯著水準 α 之下, 通常取拒絕域為 $\{F \leq c_1 \text{ 或 } F \geq c_2\}$, 其中 c_1, c_2 分別為 $\mathcal{F}_{n-1, m-1}$ 分佈之第 $\alpha/2$ 及 $1 - \alpha/2$ 分位數。

若在 H_0 之下, $\lambda(\mathbf{X})$ 的分佈知道, 且有簡單的型式, 則在給定型 I 錯誤 α 之後, 可求出 c 之值, 使(4.3)式成立。但一般而言, $\lambda(\mathbf{X})$ 往往沒有簡單的型式(如上例)。幸好當樣本數夠多時, 往往可利用下述定理來得到 $\lambda(\mathbf{X})$ 之近似分佈。其證明要用到二階泰勒展式, 在此略去。

定理4.2 設p.d.f. $f(x|\theta)$ 滿足某些平滑條件(smoothness conditions)。則在虛無假設之下，當樣本數 $n \rightarrow \infty$ 時， $-2 \log \lambda(\mathbf{X})$ 分佈收斂之 χ_k^2 分佈，其中 $k = \dim \Omega - \dim \Omega_0$ 。

在定理4.2中， $\dim \Omega$ 及 $\dim \Omega_0$ 分別表 Ω 及 Ω_0 中自由參數(free parameter)的數目。例如，在例4.1中， Ω 中有一自由的參數 μ ，故 $\dim \Omega = 1$ ；而 Ω_0 中只包含一定值 μ_0 ，故 $\dim \Omega_0 = 0$ 。事實上在此例中，

$$-2 \log \lambda(\mathbf{X}) = n(\bar{X} - \mu_0)^2 / \sigma^2$$

之分佈(對任意樣本數 n) 恰為 χ_1^2 。

8.5 一致最強力檢定

我們已建立了檢定的架構，也給了一些檢定的方法。本節我們要探討**最佳檢定**的問題。由於我們要先關切型 I 錯誤，因此所謂“最佳”就是指在顯著水準不超過一給定的值 α 之下，型 II 錯誤之機率要最小。因此在檢定 $H_0: \theta \in \Omega_0$, vs. $H_a: \theta \in \Omega_0^c$ ，我們要求對 $\forall \theta \in \Omega_0$ ，型 I 錯誤之機率皆不超過 α ，且對 $\forall \theta \in \Omega_0^c$ ，檢力函數要儘量大。我們給一定義如下。

定義5.1 檢定 $H_0: \theta \in \Omega_0$, vs. $H_a: \theta \in \Omega_0^c$ ，令 D 表所有顯著水準不超過 α 之檢定的集合。則 D 中一檢定，其檢力函數 $K(\theta)$ ，若滿足

$$(5.1) \quad K(\theta) \geq K_1(\theta), \theta \in \Omega_0^c,$$

其中 K_1 為 D 中任一檢定之檢力函數，便稱為顯著水準不超過 α 下之**一致最強力**(uniformly most powerful, 簡稱UMP)檢定，此時之拒絕域稱為顯著水準不超過 α 下之**最佳拒絕域**(best critical region)。

我們找**最佳檢定**其意義就是如定義5.1。對“顯著水準不超過 α 下之一致最強力檢定”，有時只簡短地稱為 α 下之**一致最強力檢定**(或說 α 下UMP檢定)。又若 Ω_0^c 中只有一元素(即 H_a 為簡單假設)，此時亦可將“一致”略去，只稱**最強力**(most powerful, 簡稱MP)檢定。

例5.1 設隨機變數 X 之p.d.f.為 $f(x|\theta)$, $\theta \in \Omega = \{\theta_0, \theta_1\}$ 。又設 $f(x|\theta)$ 如下表:

$\theta \setminus x$	0	1	2
θ_0	0.8	0	0.2
θ_1	0	0.8	0.2

基於 X , 欲檢定 $H_0: \theta = \theta_0$, vs. $H_a: \theta = \theta_1$ 。我們列出所有可能的拒絕域 $C_1, \dots, C_7$, 對應的顯著水準 α , 及在 θ_1 之檢力函數 $K(\theta_1)$ 如下:

$$\begin{aligned}
 C_1 &= \emptyset, \quad \alpha = 0, \quad K(\theta_1) = 0, \\
 C_2 &= \{X = 0\}, \quad \alpha = 0.8, \quad K(\theta_1) = 0, \\
 C_3 &= \{X = 1\}, \quad \alpha = 0, \quad K(\theta_1) = 0.8, \\
 C_4 &= \{X = 2\}, \quad \alpha = 0.2, \quad K(\theta_1) = 0.2, \\
 C_5 &= \{X = 0, 1\}, \quad \alpha = 0.8, \quad K(\theta_1) = 0.8, \\
 C_6 &= \{X = 0, 2\}, \quad \alpha = 1, \quad K(\theta_1) = 0.2, \\
 C_7 &= \{X = 1, 2\}, \quad \alpha = 0.2, \quad K(\theta_1) = 1, \\
 C_8 &= \{X = 0, 1, 2\}, \quad \alpha = 1, \quad K(\theta_1) = 1.
 \end{aligned}$$

在8.2節一開始我們便已說明, 一拒絕域就對應一檢定。比較8個 $K(\theta_1)$, 可看出 C_3 為 $\alpha = 0$ 下之UMP檢定($\alpha = 0$ 時, 只有此一檢定); C_7 為 $\alpha = 0.2$ 下之UMP檢定; C_7, C_8 皆為 $\alpha = 1$ 下之UMP檢定。當 $\alpha = 1$, 之所以會有兩個UMP檢定, 是因 $f(0|\theta_1) = 0$ 。實際上因 C_7 對應之 $\alpha = 0.2$, 較 C_8 對應之 $\alpha = 1$ 小, 而二者之 $K(\theta_1)$ 皆為1, 故一般是不會採拒絕域 C_8 的。

至於若 α 不為0, 0.2或1呢? 事實上對 $0 \leq \alpha < 0.2$, UMP檢定皆為 C_3 ; 對 $0.2 \leq \alpha < 1$, UMP檢定皆為 C_7 。以 $\alpha = 0.9$ 為例, 顯著水準不超過0.9之檢定有 $C_1, C_2, C_3, C_4, C_5, C_7$, 其中檢力最大的為 C_7 。並且 C_7 之顯著水準為0.2, 並沒有0.9那麼大。

在上例中, 對有些 $0 < \alpha < 1$, α 下之UMP檢定, 其顯著水準可能沒有達到 α 。連續型p.d.f.不會有這種情況, 離散型則常可能會發生。

現考慮 $f(x|\theta_1)/f(x|\theta_0)$ 。如果 $f(x|\theta_0)$ 為0, 而 $f(x|\theta_1) > 0$, 將此比值定義為 ∞ , 而如果 $f(x|\theta_1)$ 及 $f(x|\theta_0)$ 皆為0, 則定義此比值為0。則在例5.1中,

我們看到隨著 α 值由0開始增大, $f(x|\theta_1)/f(x|\theta_0)$ 愈大的 x , 依序為1,2,0, 愈先進入最強力檢定的拒絕域。這並非偶然, 底下奈曼-皮爾生引理便支持此觀察。其中虛無假設及對立假設皆設為簡單假設。

定理5.1 (奈曼-皮爾生引理). 設 n 維隨機向量 $\mathbf{X}$ 之p.d.f.為 $f(\mathbf{x}|\theta)$, $\theta \in \Omega = \{\theta_0, \theta_1\}$ 。考慮檢定 $H_0: \theta = \theta_0$, vs. $H_a: \theta = \theta_1$ 。取檢定函數滿足

$$(5.2) \quad \phi(\mathbf{x}) = \begin{cases} 1, & \text{若 } f(\mathbf{x}|\theta_1) > k f(\mathbf{x}|\theta_0), \\ 0, & \text{若 } f(\mathbf{x}|\theta_1) < k f(\mathbf{x}|\theta_0), \end{cases}$$

其中 $k \geq 0$ 為一常數。至於當 $f(\mathbf{x}|\theta_1) = k f(\mathbf{x}|\theta_0)$ 時, $\phi(\mathbf{x})$ 可任意取值0或1。令

$$(5.3) \quad \alpha = P_{\theta_0}(\phi(\mathbf{X}) = 1)。$$

則

(i) (充分性) 任一滿足(5.2)及(5.3)二式之檢定, 為一 α 下之UMP檢定。

(ii) (必要性) 若有一 $k > 0$, 使得(5.2)及(5.3)二式成立, 則每一 α 下之UMP檢定(即顯著水準不超過 α), 為一顯著水準為 α (即滿足(5.3)式)之檢定; 且每一 α 下之UMP檢定, 只除了在一滿足 $P_{\theta_0}(\mathbf{X} \in A) = P_{\theta_1}(\mathbf{X} \in A) = 0$ 之集合 A 之外, 必滿足(5.2)式。

證明.我們只證 $\mathbf{X}$ 為連續型隨機變數的情況。離散型的證法類似。

首先 $\phi(\mathbf{x})$ 若滿足(5.3)式, 便為一顯著水準 α 之檢定。此因 $\Omega_0 = \{\theta_0\}$ 只有一元素, 因此

$$\sup_{\theta \in \Omega_0} P_{\theta}(\phi(\mathbf{X}) = 1) = P_{\theta_0}(\phi(\mathbf{X}) = 1) = \alpha。$$

現令 $\phi(\mathbf{x})$ 為一滿足(5.2)及(5.3)二式之檢定函數。又令 $\phi_1(\mathbf{x})$ 為另一顯著水準不超過 α 之檢定。以

$$K(\theta) = P_{\theta}(\phi(\mathbf{X}) = 1) = E_{\theta}(\phi(\mathbf{X})) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi(\mathbf{x}) f(\mathbf{x}|\theta) d\mathbf{x},$$

$$K_1(\theta) = P_{\theta}(\phi_1(\mathbf{X}) = 1) = E_{\theta}(\phi_1(\mathbf{X})) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi_1(\mathbf{x}) f(\mathbf{x}|\theta) d\mathbf{x},$$

分別表檢定 ϕ 及 ϕ_1 之檢力函數。由於 $0 \leq \phi_1(x) \leq 1$, (5.2)式導致

$$(5.4) \quad (\phi(\mathbf{x}) - \phi_1(\mathbf{x}))(f(\mathbf{x}|\theta_1) - kf(\mathbf{x}|\theta_0)) \geq 0, \forall \mathbf{x} \in R^n。$$

因此

$$(5.5) \quad 0 \leq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (\phi(\mathbf{x}) - \phi_1(\mathbf{x}))(f(\mathbf{x}|\theta_1) - kf(\mathbf{x}|\theta_0)) d\mathbf{x} \\ = K(\theta_1) - K_1(\theta_1) - k(K(\theta_0) - K_1(\theta_0))。$$

由於 ϕ_1 之顯著水準不超過 α , 且因有(5.3)式之假設, ϕ 之顯著水準為 α , 即 $K(\theta_0) = \alpha$, 故得

$$K(\theta_0) - K_1(\theta_0) = \alpha - K_1(\theta_0) \geq 0。$$

又 $k \geq 0$, (5.5)式導致

$$0 \leq K(\theta_1) - K_1(\theta_1) - k(K(\theta_0) - K_1(\theta_0)) \leq K(\theta_1) - K_1(\theta_1),$$

即得 $K(\theta_1) \geq K_1(\theta_1)$ 。亦即 ϕ 之檢力大於 ϕ_1 之檢力。由於 ϕ_1 為任一顯著水準不超過 α 之檢定, 且 θ_1 為 Ω_0^c 中唯一的點, 故得證 ϕ 為 α 下之UMP檢定。

其次證明(ii)。設 ϕ 為一滿足(5.2)及(5.3)二式之檢定, 其中 $k > 0$, 且令 ϕ_1 為任一 α 下之UMP檢定。由(i)知 ϕ 亦為 α 下之UMP檢定, 因此二者之檢力須相等, 即

$$(5.6) \quad K(\theta_1) = K_1(\theta_1)。$$

由(5.6)式, 及 $K(\theta_0) = \alpha$ (因(5.3)式成立), 且因 $k > 0$, 檢視(5.5)式, 使得

$$(5.7) \quad \alpha - K_1(\theta_0) = K(\theta_0) - K_1(\theta_0) \leq 0。$$

現因 ϕ_1 之顯著水準不超過 α , 故 $K_1(\theta_0) \leq \alpha$ 。(5.7)式便導致 $K_1(\theta_0) = \alpha$ 。亦即 ϕ_1 為一顯著水準為 α 之檢定。即得證(ii)之前半部份。再檢視(5.5)式, 使得(5.5)式中之“ $\leq$ ”可改為“ $=$ ”。但非負函數 $(\phi(\mathbf{x}) - \phi_1(\mathbf{x}))(f(\mathbf{x}|\theta_1) - kf(\mathbf{x}|\theta_0))$ 之積分為0, 導致 ϕ_1 須滿足

$$\phi_1(\mathbf{x}) = \begin{cases} 1, & \text{若 } f(\mathbf{x}|\theta_1) > kf(\mathbf{x}|\theta_0), \\ 0, & \text{若 } f(\mathbf{x}|\theta_1) < kf(\mathbf{x}|\theta_0), \end{cases}$$

除了可能在一集合 A 不成立, 其中 A 滿足 $\int_A f(\mathbf{x}|\theta_0)dx = \int_A f(\mathbf{x}|\theta_1)dx = 0$ 。因此(ii)之後半段敘述亦得證。

註5.1 在定理5.1中, 對於 $\phi(\mathbf{x})$, 我們允許它在滿足 $f(\mathbf{x}|\theta_1) = kf(\mathbf{x}|\theta_0)$ 之集合(此集合有時不只包含一個 $\mathbf{x}$)中, 可對有些 $\mathbf{x}$ 取值為1, 有些 $\mathbf{x}$ 取值為0, 這些不同的拒絕域, 皆為UMP檢定。此處由(5.5)式亦可看出: 在 $f(\mathbf{x}|\theta_1) = kf(\mathbf{x}|\theta_0)$ 處, $\phi(\mathbf{x})$ 之值為何, 對該式並無任何影響。只是在此集合中, $\phi(\mathbf{x})$ 取值的不同, α 值(見(5.3)式)便可能不同。當然, 如果前述集合只包含一個 $\mathbf{x}$ 值, 則在該 $\mathbf{x}$ 值, $\phi(\mathbf{x})$ 只能是1或0。但若 $\mathbf{X}$ 為連續型的隨機向量, 則對任意 $0 < k < \infty$, $P(f(\mathbf{X}|\theta_1) = kf(\mathbf{X}|\theta_0)) = 0$, 所以此時不失一般性, 檢定函數可取為

$$\phi(\mathbf{x}) = \begin{cases} 1, & \text{若 } f(\mathbf{x}|\theta_1) \geq kf(\mathbf{x}|\theta_0), \\ 0, & \text{若 } f(\mathbf{x}|\theta_1) < kf(\mathbf{x}|\theta_0). \end{cases}$$

如定理5.1之上一段所提及的, 我們有時先考慮 $f(\mathbf{x}|\theta_1)/f(\mathbf{x}|\theta_0)$, 再找出 k 之值而得到如(5.2)式之檢定函數。在例5.1中, 取 $k = \infty$, 可得拒絕域 C_2 ; 取 $k = 1$, 可得拒絕域 C_5 , 取 $k = 0$, 可得拒絕域 C_7 。

例5.2 承例5.1。奈曼-皮爾生引理指出

- (i) 對 $k > 1$, 拒絕域 $= C_3 = \{X = 1\}$, 為 $-\alpha = 0$ 下之UMP檢定;
- (ii) 對 $0 < k \leq 1$, 拒絕域 $= C_7 = \{X = 1, 2\}$ 為 $-\alpha = 0.2$ 下之UMP檢定;
- (iii) 對 $k = 0$, 拒絕域 $= C_8 = \{X = 0, 1, 2\}$ 為 $-\alpha = 1$ 下之UMP檢定。

對 $\alpha = 0, 0.2, 1$, 上述三拒絕域, 與例5.1中所得相同。但這中間其實有一差異存在。在例5.1中, 對每一 $0 \leq \alpha \leq 1$, 皆能給出UMP檢定。但依奈曼-皮爾生引理, 我們只給出3個不同的 α 下之UMP檢定。你或許會以為, 在例5.1中, 對 $\alpha \in [0, 0.2)$, UMP檢定皆為 C_3 , 對 $\alpha \in [0.2, 1)$, UMP檢定皆為 C_7 , $\alpha = 1$ 下之UMP檢定則為 C_8 。所以雖在本例中, 只得到三個 α 下之UMP檢定, 但由此即可推定在 $\alpha \in [0, 0.2)$, $[0.2, 1)$, 1 , 下之UMP檢定分別是 C_3 , C_7 及 C_8 。可惜這並不正確, 見下例。

例5.3 設 X 有 $B(2, \theta)$ 分佈, 即 $f(x|\theta) = \binom{2}{x}\theta^x(1-\theta)^{2-x}$, $x = 0, 1, 2$ 。欲檢定 $H_0: \theta = 0.5$, vs. $H_a: \theta = 0.75$ 。

解.先得到

$$\begin{aligned}\frac{f(0|\theta = 0.75)}{f(0|\theta = 0.5)} &= \frac{0.25^2}{0.5^2} = 0.25, \\ \frac{f(1|\theta = 0.75)}{f(1|\theta = 0.5)} &= \frac{0.75 \cdot 0.25}{0.5^2} = 0.75, \\ \frac{f(2|\theta = 0.75)}{f(2|\theta = 0.5)} &= \frac{0.75^2}{0.5^2} = 2.25.\end{aligned}$$

奈曼-皮爾生引理指出

- (i) 對 $k > 2.25$, 拒絕域 $= \emptyset$, 為 $1-\alpha = 0$ 下之UMP檢定;
- (ii) 對 $0.75 < k \leq 2.25$, 拒絕域 $= \{X = 2\}$, 為 $1-\alpha = P(X = 2|\theta = 0.5) = 0.25$ 下之UMP檢定;
- (iii) 對 $0.25 < k \leq 0.75$, 拒絕域 $= \{X = 1, 2\}$, 為 $1-\alpha = P(X = 1 \text{ 或 } 2|\theta = 0.5) = 0.75$ 下之UMP檢定;
- (iv) 對 $0 \leq k \leq 0.25$, 拒絕域 $= \{X = 0, 1, 2\}$, 為 $1-\alpha = 1$ 下之UMP檢定。

底下我們仿例5.1, 以求出UMP檢定。先列出所有可能的拒絕域, 對應的顯著水準 α , 及在 $\theta = 0.75$ 之檢力函數 $K(0.75)$ 。

拒絕域	α	$K(0.75)$
$\emptyset$	0	0
$\{X = 0\}$	0.25	0.0625
$\{X = 1\}$	0.5	0.375
$\{X = 2\}$	0.25	0.5625
$\{X = 0, 1\}$	0.75	0.4375
$\{X = 0, 2\}$	0.5	0.625
$\{X = 1, 2\}$	0.75	0.9375
$\{X = 0, 1, 2\}$	1	1

故得各 α 下之UMP檢定如下:

α	UMP檢定中之拒絕域
$0 \leq \alpha < 0.25$	$\emptyset$
$0.25 \leq \alpha < 0.5$	$\{X = 2\}$
$0.5 \leq \alpha < 0.75$	$\{X = 0, 2\}$
$0.75 \leq \alpha < 1$	$\{X = 1, 2\}$
$\alpha = 1$	$\{X = 0, 1, 2\}$

可看出以奈曼-皮爾生引理, 得不到 $\alpha = 0.5$ 時之UMP檢定。

上例顯示, 藉由奈曼-皮爾生引理, 有時得不到所有不同 α 下之UMP檢定。離散型分佈常有此情況, 連續型分佈則通常對每一 $0 \leq \alpha \leq 1$, 都可利用奈曼-皮爾生引理得到UMP檢定。換句話說, 可由先給定的 α 值, 而求出對應的 k 值。稍後我們會給例子。另外, 對連續型的分佈, 在(5.2)式中, 若將“ $>$ ”改為“ $\geq$ ”(或將“ $<$ ”改為“ $\leq$ ”)並無影響。

例5.4 某國軍艦是由整數1開始連續編號, 總共的艘數 θ 未知, 為一正整數。欲檢定 $H_0: \theta = \theta_0$, vs. $H_a: \theta = \theta_1$, 其中 $\theta_1 > \theta_0$ 為二定值。經由隨機地觀測到的 n 艘軍艦之編號 $X_1, \dots, X_n$, 試給出UMP檢定。

解. 依假設 X 之p.d.f. 為

$$f(x|\theta) = \frac{1}{\theta}, x = 1, \dots, \theta.$$

顯然我們事先確定觀測值不會超過 θ_1 , 否則有可能 H_0 及 H_a 皆無法接受。則

$$\begin{aligned} \frac{f(\mathbf{x}|\theta_1)}{f(\mathbf{x}|\theta_0)} &= \frac{f(x_1, \dots, x_n|\theta_1)}{f(x_1, \dots, x_n|\theta_0)} \\ &= \frac{\theta_1^{-n} I_{\{x_{(n)} \leq \theta_1\}}}{\theta_0^{-n} I_{\{x_{(n)} \leq \theta_0\}}} \\ &= \begin{cases} \infty & , \text{若 } \theta_0 < x_{(n)} \leq \theta_1, \\ (\theta_0/\theta_1)^n & , \text{若 } 1 \leq x_{(n)} \leq \theta_0. \end{cases} \end{aligned}$$

取 $k = (\theta_0/\theta_1)^n$, 則滿足 $f(\mathbf{x}|\theta_1)/f(\mathbf{x}|\theta_0) = k$ 之 $\mathbf{x}$ 的集合為 $\{\mathbf{x} | 1 \leq x_{(n)} \leq \theta_0\}$ 。由奈曼-皮爾生引理, 對任意整數 $0 \leq j \leq \theta_0$, 拒絕域 $\{X_{(n)} > j\}$, 為

$1-\alpha$ 下之UMP檢定, 其中

$$\alpha = P_{\theta_0}(\text{拒絕}H_0) = P_{\theta_0}(X_{(n)} > j) = 1 - \left(\frac{j}{\theta_0}\right)^n.$$

當 j 由 0 至 θ_0 變動, 可得 $\theta_0 + 1$ 個不同的顯著水準。但如果不是如奈曼-皮爾生引理中, 允許對 $\mathbf{x} = \{\mathbf{x} | f(\mathbf{x}|\theta_1)/f(\mathbf{x}|\theta_0) = k\}$, $\phi(\mathbf{x})$ 皆可取成 0 或 1, 而是只允許 $\phi(\mathbf{x})$ 恆取為 1 或恆取為 0, 則便只有 0 及 1 兩種顯著水準:

$\phi(\mathbf{x}) = 1$, 若且唯若 $f(\mathbf{x}|\theta_1)/f(\mathbf{x}|\theta_0) \geq k$ 。

(i) 若 $k > (\theta_0/\theta_1)^n$, 則 $\alpha = P_{\theta_0}(X_{(n)} > \theta_0) = 0$;

(ii) 若 $k \leq (\theta_0/\theta_1)^n$, 則 $\alpha = P_{\theta_0}(X_{(n)} \leq \theta_0) = 1$ 。

例5.5 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{E}(\theta)$ 分佈所產生之隨機樣本, $\theta > 0$ 為已知。欲檢定 $H_0: \theta = \theta_0$, vs. $H_a: \theta = \theta_1$, 其中 $\theta_1 < \theta_0$ 為二給定的值。

由奈曼-皮爾生引理, 拒絕 H_0 , 當

$$(5.8) \quad \frac{\theta_1^n e^{-\theta_1 \sum_{i=1}^n x_i}}{\theta_0^n e^{-\theta_0 \sum_{i=1}^n x_i}} > k,$$

為一UMP檢定。給定 $0 < \alpha < 1$, k 則由

$$P_{\theta_0}\left(\frac{\theta_1^n e^{-\theta_1 \sum_{i=1}^n x_i}}{\theta_0^n e^{-\theta_0 \sum_{i=1}^n x_i}} > k\right) = \alpha$$

而決定。(5.8)式等價於

$$\sum_{i=1}^n X_i > \frac{\log(k(\theta_0/\theta_1)^n)}{\theta_0 - \theta_1} = k_1.$$

在 H_0 之下, $2\theta_0 \sum_{i=1}^n X_i \sim \chi_{2n}^2$, 故

$$2\theta_0 k_1 = \chi_{1-\alpha, 2n}^2.$$

即

$$k_1 = \frac{\chi_{1-\alpha, 2n}^2}{2\theta_0}.$$

由此得拒絕域 $\{\sum_{i=1}^n X_i > \chi_{1-\alpha, 2n}^2/(2\theta_0)\}$ 為 $1-\alpha$ 下之UMP檢定。

由 k_1 可求出 k , 不過如果只是要給出最佳拒絕域, 則並無必要求出 k 。

同理, 若欲檢定 $H_0 : \theta = \theta_0$, vs. $H_a : \theta = \theta_1$, 其中 $\theta_1 > \theta_0$, 則拒絕 H_0 , 當

$$\sum_{i=1}^n X_i < \frac{\log(k(\theta_0/\theta_1)^n)}{\theta_0 - \theta_1} = k_1。$$

故得拒絕 $\{\sum_{i=1}^n X_i < \chi_{\alpha, 2n}^2/(2\theta_0)\}$ 為 $-\alpha$ 下之UMP檢定。

對此二不同的對立假設, 因 $\theta_1 < \theta_0$ 及 $\theta_1 > \theta_0$ 之不同, 導致拒絕域不同。除此之外, 對 $H_0 : \theta = \theta_0$, vs. $H_a : \theta = \theta_1$, 及 $H_0 : \theta = \theta_0$, vs. $H_a : \theta = \theta_2$, 只要 θ_1 與 θ_2 皆大於(或小於) θ_0 , 則 α 下之UMP檢定皆相同, 與 θ_1, θ_2 之大小無關。例如, 當共同分佈為 $\mathcal{E}(\theta)$, 對於檢定

$$H_0 : \theta = 5, \text{ vs. } H_a : \theta = 6,$$

及

$$H_0 : \theta = 5, \text{ vs. } H_a : \theta = 7,$$

有相同之拒絕域。

稍後我們會看, 利用上述性質, 若對立假設為複合假設, 如 $H_a : \theta < \theta_0$, 亦可給出UMP檢定。

定理5.1亦可用來檢定母體是源自那一分佈, 見下例。

例5.6 設有一組隨機樣本 $X_1, \dots, X_n$ 。欲檢定 $H_0 : X \sim \mathcal{U}(0, 1)$, vs. $H_a : X \sim \mathcal{E}(1)$ 。

令 f_0, f_1 分別表在 H_0 及 H_a 下, $X_1, \dots, X_n$ 之聯合p.d.f.。則

$$\frac{f_1(\mathbf{x})}{f_0(\mathbf{x})} = \frac{\prod_{i=1}^n e^{-x_i}}{\prod_{i=1}^n I_{\{0 < x_i < 1\}}} = \frac{e^{-\sum_{i=1}^n x_i}}{\prod_{i=1}^n I_{\{0 < x_i < 1\}}}。$$

首先, 既然要在 H_0 及 H_a 中挑選, 表事先確定觀測值必為正, 否則有可能 H_0 及 H_a 皆無法接受。其次除非 $k = \infty$, 否則顯然若有一 $x_i \geq 1$, 就該拒絕 H_0 (此時 $f_1(\mathbf{x})/f_0(\mathbf{x}) = \infty$)。由奈曼-皮爾生引理, UMP檢定之拒絕域為 $C = \{\sum_{i=1}^n X_i \leq k_1\} \cup (\bigcup_{i=1}^n \{X_i \geq 1\})$, 其中 $k_1 = -\log k$ 。在 H_0 之下,

$\bigcup_{i=1}^n \{X_i \geq 1\}$ 發生之機率為 0, 而 $\sum_{i=1}^n X_i$ 為 n 個 i.i.d. 有 $\mathcal{U}(0, 1)$ 分佈之隨機變數的和, 其分佈並無簡單的形式。現設 n 夠大, 因此中央極限定理適用, 我們藉此來求近似的拒絕域。由於 $E(X) = 1/2$, $\text{Var}(X) = 1/12$,

$$Z_n = \frac{\bar{X} - 0.5}{\sqrt{1/(12n)}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)。$$

故

$$\{\bar{X} \leq 0.5 + z_\alpha / \sqrt{12n}\} \cup \left(\bigcup_{i=1}^n \{X_i \geq 1\} \right)$$

為一近似的顯著水準 α 之 UMP 檢定拒絕域。

底下二推論是有關充分統計量。

系理 5.1 設隨機向量 $\mathbf{X}$ 之 p.d.f. 為 $f(\mathbf{x}|\theta)$, $\theta \in \Omega = \{\theta_0, \theta_1\}$ 。考慮檢定 $H_0 : \theta = \theta_0$, vs. $H_a : \theta = \theta_1$ 。設 $T(\mathbf{X})$ 為 θ 之一充分統計量, 以 $g(t|\theta)$ 為 p.d.f.。令檢定函數滿足

$$(5.9) \quad \phi(t) = \begin{cases} 1, & \text{若 } g(t|\theta_1) > k g(t|\theta_0), \\ 0, & \text{若 } g(t|\theta_1) < k g(t|\theta_0), \end{cases}$$

其中 $k \geq 0$ 為一常數, 至於當 $g(t|\theta_1) = k g(t|\theta_0)$ 時, $\phi(t)$ 可任意取值 0 或 1。則此為一 α 下之 UMP 檢定, 其中

$$(5.10) \quad \alpha = P_{\theta_0}(\phi(T) = 1)。$$

證明. 基於 T 以 C 為拒絕域, 等價於基於 $\mathbf{X}$, 以

$$C_0 = \{T(\mathbf{X}) \in C\}$$

為拒絕域。由分解定理,

$$f(\mathbf{x}|\theta_i) = g(T(\mathbf{x})|\theta_i)h(\mathbf{x}), i = 0, 1,$$

其中 $h(\mathbf{x})$ 爲一非負函數。因此

$$\begin{aligned}\{\mathbf{x}|g(t|\theta_1)h(\mathbf{x}) > kg(t|\theta_0)h(\mathbf{x})\} &= \{\mathbf{x}|f(\mathbf{x}|\theta_1) > kf(\mathbf{x}|\theta_0)\}, \\ \{\mathbf{x}|g(t|\theta_1)h(\mathbf{x}) < kg(t|\theta_0)h(\mathbf{x})\} &= \{\mathbf{x}|f(\mathbf{x}|\theta_1) < kf(\mathbf{x}|\theta_0)\}.\end{aligned}$$

又由(5.10)式,

$$P_{\theta_0}(\mathbf{X} \in C_0) = P_{\theta_0}(T(\mathbf{X}) \in C) = \alpha.$$

故由定理5.1之充分性部分, 前述基於 T 之檢定, 爲一 α 下之UMP檢定。證畢。

我們注意到例5.4中, $X_{(n)}$ 爲 θ 之一充分統計量, 例5.5中, $\sum_{i=1}^n X_i$ 亦爲 θ 之一充分統計量, 而二例中UMP檢定之拒絕域也的確基於充分統計量。

底下爲一檢定複合假設之例子。

例5.7 設 $X_1, \dots, X_n$ 爲一組由 $\mathcal{E}(\theta)$ 分佈所產生之隨機樣本, $\theta > 0$ 爲未知。在例5.5我們已指出, 只要 $\theta_1 < \theta_0$, 對於檢定 $H_0: \theta = \theta_0$, vs. $H_a: \theta = \theta_1$, α 下之UMP檢定的拒絕域皆爲 $\{\sum_{i=1}^n X_i > \chi_{1-\alpha, 2n}^2/(2\theta_0)\}$ 。故此亦爲 $H_0: \theta = \theta_0$, vs. $H'_a: \theta < \theta_0$, α 下之UMP檢定。

此檢定之檢力函數爲

$$\begin{aligned}K(\theta) &= P_{\theta}(\text{拒絕 } H_0) \\ &= P_{\theta}\left(\sum_{i=1}^n X_i > \chi_{1-\alpha, 2n}^2/(2\theta_0)\right) \\ &= P_{\theta}\left(2\theta \sum_{i=1}^n X_i > (\theta/\theta_0)\chi_{1-\alpha, 2n}^2\right) \\ &= 1 - F_{2n}((\theta/\theta_0)\chi_{1-\alpha, 2n}^2), \quad \theta \leq \theta_0,\end{aligned}$$

其中 F_{2n} 表 χ_{2n}^2 分佈之分佈函數。此處仍用到當 θ 爲真, $2\theta \sum_{i=1}^n X_i$ 有 χ_{2n}^2 分佈。由於 $K(\theta)$ 爲 θ 之漸減函數,

$$\sup_{\theta \geq \theta_0} K(\theta) = K(\theta_0) = 1 - F_{2n}(\chi_{1-\alpha, 2n}^2) = 1 - (1 - \alpha) = \alpha.$$

故此亦為 $H'_0: \theta \geq \theta_0$, vs. $H'_a: \theta < \theta_0$, α 下之 UMP 檢定。

同理可得, 對於 $H_0: \theta = \theta_0$, vs. $H_a: \theta > \theta_0$, 或 $H''_0: \theta \leq \theta_0$, vs. $H''_a: \theta > \theta_0$, α 下 UMP 檢定之拒絕域皆為 $\{\sum_{i=1}^n X_i < \chi^2_{\alpha, 2n}/(2\theta_0)\}$, 且檢力函數為

$$K_1(\theta) = F_{2n}((\theta/\theta_0)\chi^2_{\alpha, 2n})。$$

例5.8 某電子零件之壽命可以指數分佈 $\mathcal{E}(\theta)$ 當做模式。依過去經驗, 平均壽命若小於 100 小時系統便會有問題。故欲檢定 $H_0: 1/\theta \geq 100$, vs. $H_a: 1/\theta < 100$ 。即 $H_0: \theta \leq 0.01$, vs. $H_a: \theta > 0.01$ 。

今隨機地取 40 個零件測試, 並得平均壽命 93.1 (小時)。又取 $\alpha = 0.05$ 。因 $\theta_0 = 0.01$, 由例 5.7, 拒絕域為 $\sum_{i=1}^{40} X_i < \chi^2_{0.25, 80}/(2 \cdot 0.01) = 60.39/0.02 = 3,019.5$ 。因 $\bar{X}_{40} = 93.1$, $\sum_{i=1}^{40} X_i = 3,724 > 3,019.5$ 。故在 $\alpha = 0.05$ 下, 不能拒絕 H_0 。

我們給複合假設之奈曼-皮爾生引理如下。

系理 5.2 設隨機向量 $\mathbf{X}$ 之 p.d.f. 為 $f(\mathbf{x}|\theta)$, $\theta \in \Omega = \Omega_0 \cup \Omega_0^c$ 。欲檢定 $H_0: \theta \in \Omega_0$, vs. $H_a: \theta \in \Omega_0^c$ 。設 $T(\mathbf{X})$ 為 θ 之一充分統計量, 以 $g(t|\theta)$ 為其 p.d.f.。則基於 T , 拒絕域為 C 之檢定, 若滿足下述三條件, 便為 $1-\alpha$ 下之 UMP 檢定:

- (i) 此檢定之顯著水準為 α 。
- (ii) 存在一 $\theta_0 \in \Omega_0$, 使得 $P_{\theta_0}(T \in C) = \alpha$ 。
- (iii) 對(ii)中之 θ_0 , 及對 $\forall \theta' \in \Omega_0^c$, 存在一 $k' \geq 0$, 使得若 $g(t|\theta') > k'g(t|\theta_0)$, 則 $t \in C$, 若 $g(t|\theta') < k'g(t|\theta_0)$, 則 $t \notin C$ 。

證明. 令 $K(\theta)$ 表拒絕域為 C 之檢力函數。固定一 $\theta' \in \Omega_0^c$, 考慮檢定 $H'_0: \theta = \theta_0$, vs. $H'_a: \theta = \theta'$ 。系理 5.1 及上述條件 (i)、(ii), 及 (iii), 導致對任一顯著水準不超過 α 之 H'_0 , vs. H'_a 的檢定, 其檢力函數 $K_1(\theta)$ 滿足

$$(5.11) \quad K(\theta') \geq K_1(\theta')。$$

但對任一顯著水準不超過 α 之 H_0 , vs. H_a 之檢定, 其檢力函數 $K_2(\theta)$ 滿足

$$K_2(\theta_0) \leq \sup_{\theta \in \Omega_0} K_2(\theta) \leq \alpha。$$

故此亦爲一顯著水準不超過 α 之 H'_0 , vs. H'_a 的檢定。因此由(5.11)式, $K(\theta') \geq K_2(\theta')$ 。由於 θ' 爲 Ω_0^c 中任一點, 故依定義本系理得證。

上述系理的條件(i)及(ii), 並不只是導致此乃一顯著水準爲 α 之檢定。要知對 H_0 爲複合假設, 顯著水準爲 α , 表 $\sup_{\theta \in \Omega_0} K(\theta) = \alpha$, 其中 $K(\theta)$ 爲檢力函數。但有可能不存在一 $\theta \in \Omega_0$, 使得 $K(\theta) = \alpha$ 。例如, 設 $K(\theta) = \theta$, $\Omega_0 = [0, 0.5]$, 則 $\sup_{\theta \in \Omega_0} K(\theta) = 0.5$ 。但對 $\forall \theta \in \Omega_0$, $K(\theta)$ 皆小於0.5, 並不存在一 $\theta \in \Omega_0$, 使得 $K(\theta) = 0.5$ 。條件(i)及(ii)要求在 Ω_0 中, $K(\theta)$ 要能達到最小上界。換句話說, 最小上界此時成爲極大值(maximum)。

讀者可試以系理5.2重做例2.7。

例5.9 設隨機變數 X 之p.d.f. $f(x|\theta)$, $\theta \in \Omega = \{\theta_0, \theta_1, \theta_2, \theta_3\}$, 如表5.1所給。欲檢定 $H_0: \theta = \theta_0$, vs. $H_a: \theta \in \{\theta_1, \theta_2, \theta_3\}$ 。則是否存在一 $\alpha = 2/5$ 下之UMP檢定?

表5.1 例5.9中之p.d.f. $f(x|\theta)$

$\theta \setminus x$	1	2	3	4	5	6
θ_0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	0
θ_1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{16}$
θ_2	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{8}$
θ_3	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$

解. 令 $\Omega_0 = \{\theta_0\}$, $\Omega_0^c = \{\theta_1, \theta_2, \theta_3\}$ 。考慮

$$\frac{f(x|\theta)}{f(x|\theta_0)}, \theta \in \Omega_0^c,$$

並取檢定函數爲

$$\phi(x) = \begin{cases} 1, & \text{若 } f(x|\theta)/f(x|\theta_0) \geq 5/8, \forall \theta \in \Omega_0^c, \\ 0, & \text{其他,} \end{cases}$$

$$= \begin{cases} 1, & \text{若 } x \in \{3, 5, 6\}, \\ 0, & \text{若 } x \in \{1, 2, 4\}. \end{cases}$$

此檢定之顯著水準

$$\alpha = f(3|\theta_0) + f(5|\theta_0) + f(6|\theta_0) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}.$$

因此由系理5.2知, 此為 $1-\alpha = 3/5$ 下之UMP檢定。

如果不利用系理5.2, 由於這是一離散型的p.d.f., 且 x 及 θ 之可能取的值也不太多, 因此可仿例5.1之“窮舉法”。即將所有可能的拒絕域均列出, 再比較 $\theta \in \Omega_0^c$ 時, 各檢力函數之大小。我們推導如下。

由於 $\alpha = 2/5$, 拒絕域 C 要包含1,2,3,4,5中的兩個數, 如此型 I 錯誤之機率才會 $= 1/5 + 1/5 = 2/5$ 。又因要使 $\theta \in \Omega_0^c$ 時, 檢力愈大愈好, 而 $f(6|\theta_0) = 0$, 所以拒絕域要納入6這一數(為什麼?)。即拒絕域要有1,2,3,4,5中的兩個數及6, 共有 $\binom{5}{2} = 10$ 種可能。依上討論, 表5.2給出所有可能的拒絕域及對應的 $\theta \in \Omega_0^c$ 時之檢力函數。

表5.2 例5.9中 $\theta \in \Omega_0^c$ 之檢力函數

$C \setminus \theta$	θ_1	θ_2	θ_3
$\{X = 1, 2, 6\}$	$\frac{4}{16}$	$\frac{8}{16}$	$\frac{10}{16}$
$\{X = 1, 3, 6\}$	$\frac{7}{16}$	$\frac{8}{16}$	$\frac{10}{16}$
$\{X = 1, 4, 6\}$	$\frac{7}{16}$	$\frac{4}{16}$	$\frac{10}{16}$
$\{X = 1, 5, 6\}$	$\frac{7}{16}$	$\frac{8}{16}$	$\frac{10}{16}$
$\{X = 2, 3, 6\}$	$\frac{6}{16}$	$\frac{10}{16}$	$\frac{12}{16}$
$\{X = 2, 4, 6\}$	$\frac{6}{16}$	$\frac{6}{16}$	$\frac{12}{16}$
$\{X = 2, 5, 6\}$	$\frac{6}{16}$	$\frac{10}{16}$	$\frac{12}{16}$
$\{X = 3, 4, 6\}$	$\frac{9}{16}$	$\frac{6}{16}$	$\frac{12}{16}$
$\{X = 3, 5, 6\}$	$\frac{9}{16}$	$\frac{10}{16}$	$\frac{12}{16}$
$\{X = 4, 5, 6\}$	$\frac{9}{16}$	$\frac{6}{16}$	$\frac{12}{16}$

由表5.2, 對拒絕域為 $C = \{X = 3, 5, 6\}$, $\theta \in \Omega_0^c$ 之檢力函數為

$$K(\theta_1) = \frac{9}{16}, \quad K(\theta_2) = \frac{10}{16}, \quad K(\theta_3) = \frac{12}{16},$$

顯然滿足 $K(\theta_i) \geq K_1(\theta_i)$, $i = 1, 2, 3$, 其中 K_1 為表 5.2 中任一檢定之檢力函數。故取拒絕域 $\{X = 3, 5, 6\}$, 為 $\alpha = 2/5$ 下之 UMP 檢定。

例 5.10 設 X 有 $\mathcal{B}(n, \theta)$ 分佈, n 為已知。欲檢定 $H_0 : \theta \leq \theta_0$, vs. $H_a : \theta > \theta_0$ 。在例 2.6 中(在該例裡 $n = 5$, $\theta_0 = 0.5$), 我們以直觀的方式取拒絕域 $\{X \geq k\}$, 且顯著水準為

$$\alpha = \sup_{\theta \leq \theta_0} P_\theta(X \geq k) = P_{\theta_0}(X \geq k) = \sum_{j=k}^n \binom{n}{j} \theta_0^j (1 - \theta_0)^{n-j}。$$

對 $\mathcal{B}(n, \theta)$ 分佈, $P_\theta(X \geq k)$ 為 $-\theta$ 之漸減函數, 證明留在習題第 21 題。底下我們來驗證此確為 $1-\alpha$ 下之 UMP 檢定。

對任意 $\theta_1 > \theta_0$,

$$\begin{aligned} \frac{f(x|\theta_1)}{f(x|\theta_0)} &= \frac{\binom{n}{x} \theta_1^x (1 - \theta_1)^{n-x}}{\binom{n}{x} \theta_0^x (1 - \theta_0)^{n-x}} \\ &= \left(\frac{\theta_1/(1 - \theta_1)}{\theta_0/(1 - \theta_0)} \right)^x \left(\frac{1 - \theta_1}{1 - \theta_0} \right)^n \\ &= \rho^x \left(\frac{1 - \theta_1}{1 - \theta_0} \right)^n, \end{aligned}$$

其中

$$\rho = \frac{\theta_1/(1 - \theta_1)}{\theta_0/(1 - \theta_0)} = \frac{\theta_1(1 - \theta_0)}{\theta_0(1 - \theta_1)}。$$

因 $\theta_1 > \theta_0$, 故 $\rho > 1$ 。因此

$$\frac{f(x|\theta_1)}{f(x|\theta_0)} \geq k \Leftrightarrow x \geq \frac{\log k + n \log((1 - \theta_0)/(1 - \theta_1))}{\log \rho} = c。$$

故由系理 5.2, 取拒絕域 $C = \{X \geq c\}$, 再選 c , 使顯著水準為 α , 便為 UMP 檢定, 即 c 滿足 $\alpha = P_{\theta_0}(X \geq c)$ 。

附帶一提, 本例若改為簡單假設, 即設 $H_0 : \theta = \theta_0$, vs. $H_a : \theta = \theta_1$, 其中 $\theta_1 > \theta_0$, 則仍得到相同型式之 UMP 檢定的拒絕域 $\{X \geq c\}$ 。

例 5.11 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(0, \theta)$ 分佈所產生之隨機樣本, 其中 $\theta > 0$ 。對檢定 $H_0 : \theta = \theta_0$, vs. $H_a : \theta > \theta_0$, 試求 $1-\alpha$ 下之 UMP 檢定, 其中 θ_0 為

一固定正數。

解. 首先

$$f(\mathbf{x}|\theta) = f(x_1, \dots, x_n|\theta) = \left(\frac{1}{2\pi\theta}\right)^{n/2} e^{-\sum_{i=1}^n x_i^2/(2\theta)}。$$

對 $\theta_1 > \theta_0$,

$$\begin{aligned} \frac{f(\mathbf{x}|\theta_1)}{f(\mathbf{x}|\theta_0)} &= \left(\frac{\theta_0}{\theta_1}\right)^{n/2} \exp\left\{-\frac{\theta_0 - \theta_1}{2\theta_0\theta_1} \sum_{i=1}^n x_i^2\right\} > k \\ \Leftrightarrow \sum_{i=1}^n x_i^2 &> \frac{2\theta_0\theta_1}{\theta_1 - \theta_0} \left(\frac{n}{2} \log \frac{\theta_1}{\theta_0} + \log k\right) = c。 \end{aligned}$$

故由系理5.2, 取拒絕域為 $C = \{\sum_{i=1}^n X_i^2 > c\}$, 再選 c 使得顯著水準為 α , 便為UMP檢定。即

$$\alpha = P_{\theta_0}\left(\sum_{i=1}^n X_i^2 > c\right) = P_{\theta_0}\left(\sum_{i=1}^n X_i^2/\theta_0 > c/\theta_0\right)。$$

由於在 H_0 之下, $\sum_{i=1}^n X_i^2/\theta_0$ 有 χ_n^2 分佈, 故由

$$\chi_{1-\alpha, n}^2 = \frac{c}{\theta_0},$$

解出 $c = \theta_0 \chi_{1-\alpha, n}^2$ 。

例如, 取 $n = 15$, $\alpha = 0.05$, 且 $\theta_0 = 3$ 。由附表5, 因 $\chi_{0.95, 15}^2 \doteq 25.00$, 故

$$c \doteq 3 \cdot 25 = 75。$$

即拒絕域 $\{\sum_{i=1}^n X_i^2 > 75\}$, 為 $\alpha = 0.05$ 下之UMP檢定。

例5.12 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\theta, \sigma^2)$ 分佈所產生之隨機樣本, σ^2 設為已知。則 $\bar{X}$ 為 θ 之一充分統計量。

先檢定 $H_0: \theta = \theta_0$, vs. $H_a: \theta = \theta_1$, 其中 $\theta_1 < \theta_0$ 。首先 $\bar{X}$ 之 p.d.f. 為

$$g(\bar{x}|\theta) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-n(\bar{x}-\theta)^2/(2\sigma^2)}, \bar{x} \in R,$$

$g(\bar{x}|\theta_1) > kg(\bar{x}|\theta_0)$ 等價於

$$\begin{aligned} & \exp\{-n((\bar{x} - \theta_1)^2 - (\bar{x} - \theta_0)^2)/(2\sigma^2)\} > k \\ \Leftrightarrow & -(2\bar{x}(\theta_0 - \theta_1) + (\theta_1^2 - \theta_0^2)) > (2\sigma^2 \log k)/n \\ \Leftrightarrow & 2\bar{x}(\theta_0 - \theta_1) < -(2\sigma^2 \log k)/n - (\theta_1^2 - \theta_0^2) \\ \Leftrightarrow & \bar{x} < \frac{(2\sigma^2 \log k)/n + \theta_1^2 - \theta_0^2}{2(\theta_1 - \theta_0)}. \end{aligned}$$

此處用到 $\theta_1 < \theta_0$ 。當 k 由 0 漸增至 ∞ ，上述最後一不等式的右側，由 ∞ 漸減至 $-\infty$ 。由系理 5.1，以 $\{\bar{X} < c\}$ 為拒絕域的檢定，為 $-\alpha$ 下之 UMP 檢定，其中 $\alpha = P_{\theta_0}(\bar{X} < c)$ 。給定 $-\alpha$ 值，UMP 檢定之拒絕域為 $\{\bar{X} < c = \theta_0 + z_\alpha \sigma / \sqrt{n}\}$ 。

其次考慮檢定 $H'_0 : \theta \geq \theta_0$, vs. $H'_a : \theta < \theta_0$ 。底下說明拒絕域 $\{\bar{X} < \theta_0 + z_\alpha \sigma / \sqrt{n}\}$ ，仍為 $-\alpha$ 下之 UMP 檢定。

系理 5.2 中之條件 (ii) 當然成立 (取前述 θ_0 當做 (ii) 中的 θ_0 即可)。又條件 (iii) 成立，此因在前述推導 $H_0 : \theta = \theta_0$, vs. $H_a : \theta = \theta_1$ 的 α 下之 UMP 檢定中，只須 $\theta_1 < \theta_0$ ，而不須知道是那一特定的 θ_1 。條件 (i) 也成立，因為此檢定之檢力函數

$$\begin{aligned} K(\theta) &= P_\theta(\bar{X} < z_\alpha \sigma / \sqrt{n} + \theta_0) \\ &= P_\theta\left(\frac{\bar{X} - \theta}{\sigma / \sqrt{n}} < z_\alpha + \frac{\theta_0 - \theta}{\sigma / \sqrt{n}}\right), \end{aligned}$$

為 $-\theta$ 之漸減函數，故

$$\sup_{\theta \geq \theta_0} K(\theta) = K(\theta_0) = \alpha。$$

由於系理 5.2 之條件 (i)、(ii) 及 (iii) 皆成立，故得證拒絕域 $\{\bar{X} < \theta_0 + z_\alpha \sigma / \sqrt{n}\}$ ，為 $-\alpha$ 下之 UMP 檢定。

最後，對 $H''_0 : \theta \leq \theta_0$, vs. $H''_a : \theta > \theta_0$ ，仿前述情況之推導 (見習題第 26 題)，可得拒絕域 $\{\bar{X} > \theta_0 + z_{1-\alpha} \sigma / \sqrt{n}\}$ ，為 $-\alpha$ 下之 UMP 檢定。

UMP 檢定有時並不存在，見下例。

例5.13 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\theta, \sigma^2)$ 分佈所產生之隨機樣本, σ^2 設為已知。欲檢定 $H_0: \theta = \theta_0$, vs. $H_a: \theta \neq \theta_0$ 。

對 $\theta_1 \neq \theta_0$, 由上例, $g(\bar{x}|\theta_1) > kg(\bar{x}|\theta_0)$ 等價於

$$2\bar{x}(\theta_0 - \theta_1) < -(2\sigma^2 \log k)/n - (\theta_1^2 - \theta_0^2) \\ \Leftrightarrow \begin{cases} \bar{x} < \frac{(2\sigma^2 \log k)/n + \theta_1^2 - \theta_0^2}{2(\theta_1 - \theta_0)}, & \text{若 } \theta_1 < \theta_0, \\ \bar{x} > \frac{(2\sigma^2 \log k)/n + \theta_1^2 - \theta_0^2}{2(\theta_1 - \theta_0)}, & \text{若 } \theta_1 > \theta_0. \end{cases}$$

故在不同的 H_a 下, 例如, 對 $H'_a: \theta = \theta_0 - 1$, 與 $H''_a: \theta = \theta_0 + 1$, 最佳拒絕域不同。因此由定義5.1, UMP檢定不存在。

事實上, 若 $H_0: \theta \in \Omega_0 = \{\theta_0\}$ 為簡單假設, $H_a: \theta \in \Omega_0^c$ 為複合假設, 則 $1-\alpha$ 下之 UMP 檢定, 必亦為 $H_0: \theta = \theta_0$, vs. $H'_a: \theta = \theta_1$ 的 α 下之 UMP 檢定 (即有相同之最佳拒絕域), 其中 $\theta_1 \in \Omega_0^c$ 。上例就是利用此性質, 以判定 UMP 檢定不存在。

對單邊假設, 有一類分佈, 其 UMP 檢定必會存在。我們先給下述定義:

定義5.2 設隨機變數 X 之 p.d.f. 為 $f(x|\theta)$, $\theta \in \Omega$, 其中 Ω 為一實數的子集合。令 $T = T(X)$ 。若對 $\forall \theta_2 > \theta_1$, $f(x|\theta_2)/f(x|\theta_1)$ 為 $T(x)$ 之非漸減函數, 其中 $x \in \{x|f(x|\theta_1) > 0, \text{ 或 } f(x|\theta_2) > 0\}$, 則稱分佈族 $\{f(x|\theta), \theta \in \Omega\}$ 對 $T = T(X)$ 有 **單調概似比** (monotone likelihood ratio, 簡寫 MLR)。

在上述定義中, X 亦可為一隨機向量。又為了便利, 若 $c > 0$, 我們將 $c/0$ 定義為 ∞ 。

許多常見的分佈族皆有 MLR。例如, 常態分佈 (假設變異數已知, 期望值未知), 波松分佈及二項分佈等。以波松分佈為例。設 X 有 $\mathcal{P}(\theta)$ 分佈, p.d.f. 為

$$f(x|\theta) = \frac{e^{-\theta} \theta^x}{x!}, x = 0, 1, \dots。$$

對 $\theta_2 > \theta_1$, 因

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{e^{-\theta_2} \theta_2^x / x!}{e^{-\theta_1} \theta_1^x / x!} = e^{-(\theta_2 - \theta_1)} \left(\frac{\theta_2}{\theta_1}\right)^x$$

為 x 之非漸減函數(事實上為嚴格漸增), 故分佈族 $\{\mathcal{P}(\theta), \theta > 0\}$ 有MLR。我們給一定理如下, 其證明不難, 因此略去。

定理5.2 設 X 之p.d.f.為

$$(5.12) \quad f(x|\theta) = h(x)c(\theta)\exp\{w(\theta)t(x)\}I_A(x), \theta \in \Omega。$$

且設 $w(\theta)$ 為 θ 之非漸減函數。即 X 有1個參數(θ 為一維)之指數族分佈(定義見5.2節), 則充分統計量 $T = T(X)$ 之分佈族 $\{g(t|\theta), \theta \in \Omega\}$ 有MLR。

底下就是保證對MLR分佈族, UMP檢定必會存在的定理。

定理5.3 考慮檢定 $H_0 : \theta \leq \theta_0$, vs. $H_a : \theta > \theta_0$ 。設 T 為 θ 之一充分統計量, 且 T 之分佈族 $\{g(t|\theta), \theta \in \Omega\}$ 有MLR。則對任意 t_0 , 拒絕域 $\{T > t_0\}$, 為 $-\alpha$ 下之UMP檢定, 其中 $\alpha = P_{\theta_0}(T > t_0)$ 。

證明. 首先由於 T 之分佈族有MLR, 故檢力函數

$$K(\theta) = P_{\theta}(T > t_0)$$

為非漸增(習題第33題)。即得

$$\sup_{\theta \leq \theta_0} K(\theta) = K(\theta_0) = \alpha,$$

故此檢定之顯著水準為 α 。又由假設知, 系理5.2之條件(ii)成立。最後對 $\theta' > \theta_0$, 若取

$$k' = \inf_{t \in J} \frac{g(t|\theta')}{g(t|\theta_0)},$$

其中 $J = \{t|t > t_0, \text{ 且 } g(t|\theta') > 0 \text{ 或 } g(t|\theta_0) > 0\}$, 則系理5.2之條件(iii)成立。故由系理5.2, 此為 $-\alpha$ 下之UMP檢定。

同理可證, 在定理5.3之條件下, 若欲檢定 $H_0 : \theta \geq \theta_0$, vs. $H_a : \theta < \theta_0$, 則拒絕域 $\{T < t_0\}$, 為 $-\alpha$ 下之UMP檢定, 其中 $\alpha = P_{\theta_0}(T < t_0)$ 。

利用定理5.3, 有些UMP檢定, 便可較容易地得到。

例5.14 承例5.12。因 $T = \bar{X}$ 有 $\mathcal{N}(\theta, \sigma^2/n)$ 分佈, 故 T 之p.d.f.為

$$g(t|\theta) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{nt^2}{2\sigma^2} - \frac{n\theta^2}{2\sigma^2} + \frac{n\theta}{\sigma^2}t\right\}, \theta \in R。$$

因 $w(\theta) = n\theta/\sigma^2$ 為 θ 之漸增函數, 故由定理5.2, $\{g(t|\theta), \theta \in R\}$ 有MLR。利用定理5.3, 即得對檢定 $H_0: \theta \leq \theta_0$, vs. $H_a: \theta > \theta_0$, 拒絕域 $\{T > t_0\}$, 為 $-\alpha$ 下之UMP檢定, 其中 $\alpha = P_{\theta_0}(T > t_0)$ 。由此立即又得 $t_0 = \theta_0 + z_{1-\alpha}\sigma/\sqrt{n}$, 與例5.12所得相同。

例5.15 承例5.10。因

$$f(x|\theta) = \binom{n}{x} (1-\theta)^n \exp\{x \log(\theta/(1-\theta))\},$$

且

$$c(\theta) = \log \frac{\theta}{1-\theta} = \log\left(\frac{1}{1-\theta} - 1\right)$$

為 θ 之漸增函數, 故由定理5.3, 即得對檢定 $H_0: \theta \leq \theta_0$, vs. $H_a: \theta > \theta_0$, 拒絕域 $\{T > t_0\}$, 為 $-\alpha$ 下之UMP檢定, 其中 $\alpha = P_{\theta_0}(T > t_0)$, 與例5.10所得相同。

例5.16 設 $X_1, \dots, X_n$ 為一組由 $\Gamma(p, \theta)$ 分佈所產生之隨機樣本, p 設為已知, $\theta > 0$ 。試對下述二檢定

(i) $H_0: \theta \leq \theta_0$, vs. $H_a: \theta > \theta_0$,

(ii) $H_0: \theta \geq \theta_0$, vs. $H_a: \theta < \theta_0$,

各給出 α 下之UMP檢定。

解. 首先 $X_1, \dots, X_n$ 之聯合p.d.f.為

$$\begin{aligned} f(\mathbf{x}|\theta) &= \prod_{i=1}^n \frac{x_i^{p-1} e^{-x_i/\theta}}{\Gamma(p)\theta^p} \\ &= (\Gamma(p)\theta^p)^{-n} \exp\left\{(p-1) \sum_{i=1}^n \log x_i - \frac{1}{\theta} \sum_{i=1}^n x_i\right\}, x_i > 0, i = 1, \dots, n。 \end{aligned}$$

(i) 因 $w(\theta) = -1/\theta$ 為 θ 之漸增函數, 故充分統計量 $T = T(\mathbf{X}) = \sum_{i=1}^n X_i$ 之分佈族有 MLR。由定理 5.3, 即得對檢定 $H_0: \theta \leq \theta_0$, vs. $H_a: \theta > \theta_0$, 拒絕域 $\{T = \sum_{i=1}^n X_i > t_0\}$, 為 $1-\alpha$ 下之 UMP 檢定, 其中 $\alpha = P_{\theta_0}(\sum_{i=1}^n X_i > t_0)$ 。因 $\sum_{i=1}^n X_i$ 有 $\Gamma(np, \theta)$ 分佈, 故當 $\theta = \theta_0$, $\sum_{i=1}^n X_i/\theta_0$ 有 $\Gamma(np, 1)$ 分佈。因此 $t_0 = \theta_0 g_{1-\alpha, np}$, 其中若 U 有 $\Gamma(np, 1)$ 分佈, 則對 $\forall 0 < u < 1$,

$$g_{u, np} = P(U \leq u)。$$

又此時之檢力函數為

$$K(\theta) = P_{\theta}(\sum_{i=1}^n X_i \geq t_0) = P_{\theta}(\frac{1}{\theta} \sum_{i=1}^n X_i \geq \frac{t_0}{\theta}) = 1 - G_{np}(\frac{\theta_0}{\theta} g_{1-\alpha, np}),$$

$\theta > 0$, 其中 G_{np} 表 $\Gamma(np, 1)$ 之分佈函數。顯然 $K(\theta)$ 為 θ 之漸增函數。

(ii) 令 $\eta = 1/\theta$, $\eta_0 = 1/\theta_0$ 。則 $0 < \eta < \infty$, 且

$$\begin{aligned} f(\mathbf{x}|\eta) &= \prod_{i=1}^n \frac{\eta^p x_i^{p-1} e^{-\eta x_i}}{\Gamma(p)} \\ &= (\eta^p / \Gamma(p))^n \exp\{(p-1) \sum_{i=1}^n \log x_i - \eta \sum_{i=1}^n x_i\}, x_i > 0, i = 1, \dots, n。 \end{aligned}$$

又此時檢定成為 $H_0: \eta \leq \eta_0$, vs. $H_a: \eta > \eta_0$ 。

因 $w(\eta) = \eta$ 為 η 之漸增函數, 故充分統計量 $T_1 = T_1(\mathbf{X}) = -\sum_{i=1}^n X_i$ 之分佈族有 MLR。由定理 5.3, 即得對檢定 $H_0: \eta \leq \eta_0$, vs. $H_a: \eta > \eta_0$, 拒絕域 $\{T_1 = -\sum_{i=1}^n X_i > t_0\}$, 為一 UMP 檢定, 且顯著水準

$$\alpha = P_{\theta_0}(-\sum_{i=1}^n X_i > t_0) = P_{\theta_0}(\sum_{i=1}^n X_i < -t_0)。$$

仿(i), 可得 $-t_0/\theta_0 = g_{\alpha, np}$, 即 $t_0 = -\theta_0 g_{\alpha, np}$ 。檢力函數則為

$$\begin{aligned} K_1(\theta) &= P_{\theta}(-\sum_{i=1}^n X_i > t_0) = P_{\theta}(\frac{1}{\theta} \sum_{i=1}^n X_i < -t_0) \\ &= P_{\theta}(\frac{1}{\theta} \sum_{i=1}^n X_i < \frac{\theta_0}{\theta} g_{\alpha, np}) \\ &= G_{np}(\frac{\theta_0}{\theta} g_{\alpha, np}), \theta > 0, \end{aligned}$$

為 θ 之漸減函數。

例5.17 設 $X_1, \dots, X_n$ 為一組由p.d.f. $f(x|\theta) = e^{-(x-\theta)}$, $x \geq \theta$, 所產生之隨機樣本, 其中 $\theta \in R$ 。欲檢定 $H_0: \theta \leq \theta_0$, vs. $H_a: \theta > \theta_0$ 。

首先注意 X 並無指數族分佈。對 $\forall \theta_2 > \theta_1$,

$$\begin{aligned} \frac{f(\mathbf{x}|\theta_2)}{f(\mathbf{x}|\theta_1)} &= \frac{\prod_{i=1}^n (e^{-(x_i-\theta_2)} I_{\{x_i \geq \theta_2\}})}{\prod_{i=1}^n (e^{-(x_i-\theta_1)} I_{\{x_i \geq \theta_1\}})} \\ &= \frac{e^{n\theta_2} I_{\{x_{(1)} \geq \theta_2\}}}{e^{n\theta_1} I_{\{x_{(1)} \geq \theta_1\}}} \\ &= \begin{cases} 0 & , \theta_1 \leq x_{(1)} < \theta_2, \\ e^{n(\theta_2-\theta_1)} & , x_{(1)} \geq \theta_2. \end{cases} \end{aligned}$$

上述比值在 $x_{(1)} < \theta_1$ 無定義, 但此並非問題, 因 $P_\theta(X_{(1)} < \theta_1) = 0$ 。由於上述比值為 $x_{(1)}$ 之非漸減函數, 故分佈族 $\{f(\mathbf{x}|\theta), \theta \in R\}$ 對充分統計量 $T = X_{(1)}$ 有MLR。由定理5.3, α 下之UMP檢定的拒絕域為 $\{X_{(1)} > k\}$, 其中

$$\alpha = P_{\theta_0}(X_{(1)} > k) = e^{-n(k-\theta_0)}。$$

故 $k = \theta_0 - (\log \alpha)/n$ 。此UMP檢定之拒絕域, 與例4.2中利用LRT檢定所得之拒絕域相同。

8.6 結語

胡適名言“大膽的假設, 小心的求證”, 是不少人的座右銘。這句話曾在學術界引起爭論, 反對者有人主張宜改為“大膽的懷疑, 小心的求證”。在假設檢定裡, 假設的產生(尤指對立假設), 乃來自觀察所得之證據, 絕非只是大膽而已。若能善用統計資訊, 則便能更科學地採合理懷疑的態度, 而非只是大膽的懷疑。因此統計裡的假設檢定, 秉持的乃是“合理的懷疑, 合理的假設, 小心的求證”之精神。

此外, 要做一有關銅板出現正面機率之檢定, 其程序通常就是連續投擲銅板若干次, 此過程一般是較單純的。但若要檢定抽煙是否致癌, 可能就不是那麼容易了。要找一群人來做實驗, 而且得長期的觀察。每個人體

質、個性、工作環境不盡相同，抽煙的量及習性也不同。實驗中途，有人說不定因某種緣故退出。而不抽煙說不定也有得肺癌者。即使想對本章一開始那則女士喝茶的情況做一檢定，拿出10個杯子，依序一杯先放茶，一杯先放牛奶，這樣妥當嗎？顯然不太好，說不定她喝了幾杯後，便猜出放茶與牛奶的順序是有規律了。因此在實驗一開始，要有好的設計，如此得到的結論才會可靠。統計學裡的實驗設計，便是討論這類問題。這是大統計學家費雪發展出來的領域。

一種新藥問世，如何設計一實驗以檢定其藥效呢？最基本的方法就是比較。要有兩組病人，一組稱為**處理組**(treatment group)，一組稱為**對照組**(control group)，然後比較此二組病人服用新藥後的反應。如何將病人分組呢？通常以隨機的方式，以排除諸如較健康者集中在同一組的偏差情況。又實驗必須在**雙盲**(double blind)下進行：測定反應的醫生及參與的病人皆不知道誰在處理組，誰在對照組。

大家再想想，8.1節所提到的古埃及人檢驗女性懷孕的方法，應可理解那不僅是假設檢定的問題而已。實驗設計是統計學裡的**紙上談兵**。做一實驗前，先要籌劃，留意統計裡隨機性的特質，以及不同個體間的差異。要知並非每個個體，都可視為一個樂意服從伯努力分佈之變化的銅板來處理。

對實驗設計有興趣者，可參考羅夢娜(1987a, 1987b)二文。

習 題

第8.2節

1. 設 X 有 $\mathcal{N}(\mu, 16)$ 分佈。欲檢定 $H_0: \mu = 10$, vs. $H_a: \mu = 11$ ，取樣本數 $n = 25$ 之一組隨機樣本。試決定型 I 錯誤機率 $\alpha = 0.05$ 下的一拒絕域，並求此時之型 II 錯誤之機率。(解. $\{\bar{X} > 11.316\}$, 0.654)
2. 在上題中，試證 $\{10 < \bar{X} < 10.1006\}$ 亦為一型 I 錯誤機率為 $\alpha = 0.05$ 之拒絕域，並求此時之型 II 錯誤之機率。(解. 0.9752)
3. 在第1題中，設 $n = 100$ ，其餘不變。試求此時之一拒絕域，及型 II 錯

誤之機率。(解. $\{\bar{X} > 10.658\}$, 0.196)

4. 在第1題中, 設拒絕域為 $\{\bar{X} > c\}$ 。試求 $\bar{X} = 11.40$ 時之 p -值。(解. 0.0401)
5. 設 $X_1, \dots, X_{36}$ 為一組由 $\mathcal{N}(\theta, 36)$ 分佈所產生之隨機樣本。欲檢定 $H_0: \theta = 10$, vs. $H_a: \theta = 12$ 。拒絕域取為 $\{\bar{X} > 11\}$ 。試決定型 I 及型 II 錯誤之機率 α, β 。(解. $\alpha = \beta \doteq 0.1587$)
6. 設 $X_1, \dots, X_n$ 為一組由 $\text{Ber}(\theta)$ 分佈所產生之隨機樣本。欲檢定 $H_0: \theta = 1/4$, vs. $H_a: \theta > 1/2$ 。取拒絕域為 $\{\sum_{i=1}^n X_i > c\}$ 。當 $n = 20$, $c = 8$, 試決定型 I 錯誤之機率, 及 $\theta = 0.5, 0.7$ 時型 II 錯誤之機率。(解. 0.0409, 0.2517, 0.0051)
7. 在上題中, 設 $n = 20$, $c = 7$ 。試決定型 I 錯誤之機率, 及 $\theta = 0.5$ 時型 II 錯誤之機率。(解. 0.1018, 0.1316)
8. 在第6題中, 設 $n = 100$, $c = 36$ 。試決定近似的型 I 錯誤之機率, 及 $\theta = 0.5$ 時型 II 錯誤之機率。(解. 0.0039, 0.0035)
9. 在例2.2中, 取 $c = 55$ 。試求此時之 α 及 $\beta(0.6)$ 之值。(解. $\alpha \doteq 0.1841$, $\beta(0.6) \doteq 0.1308$)
10. 試證例2.6中之 $K_1(\theta)$, $K_2(\theta)$ 皆為 θ 之漸增函數。
11. 試繪例2.7中 $y = K(\mu)$ 之圖形, $\mu \in R$, 其中 c, σ, n 設為常數, 給出漸近線、極值及反曲點。並驗證 $y = K(\mu)$ 為一嚴格漸增函數。(解. 水平漸近線 $y = 0, y = 1$, 無極值, 反曲點在 $\mu = c$)
12. 在例2.7中, 欲 $\alpha = 0.05$, 且當 $\mu \geq \mu_0 + \sigma$ 時, 型 II 錯誤之機率不超過0.1。
 - (i) 試問如何選取 c (表示成 μ_0 及 σ 之函數)及樣本數 n ;
 - (ii) 對(i)中的 c 及 n , 試求 $\mu \geq \mu_0 + \sigma$ 時, 型 II 錯誤機率之最小上界(求近似值至小數第四位)。
 (解. (i) n 至少為9, $n = 9$ 時, $c = \mu_0 + 1.645\sigma/3$, (ii) 0.0877)

13. 試證在例2.11中, $K_1(\theta)$ 及 $K_2(\theta)$ 皆不滿足(2.17)式。
14. 在例2.11中, 試說明基於 $X_{(n)}$ 之檢定為偏差的。次設拒絕域為 $C = \{X_{(n)} < a, \text{或} X_{(n)} > 5.9\}$ 。試求 a 之值, 使此檢定成為不偏的。(解. $a^{15} = (1 - (5.9/6)^{15}) / ((1/5)^{15} - (1/6)^{15})$, $a \doteq 4.544076$)
15. 設 X_1, X_2 為由 $Be(\theta, 1)$ 分佈所產生之隨機樣本, $\theta \in \Omega = \{1, 2\}$ 。欲檢定 $H_0 : \theta = 1$, vs. $H_a : \theta = 2$ 。拒絕域取為 $\{(3/4)X_1 \leq X_2\}$ 。試求此檢定之檢力函數。(解. $K(1) = 1/4 + (3/4)\log(3/4)$, $K(2) = 7/16 + (9/8)\log(3/4)$)
16. 設 X 有 $B(10, \theta)$ 分佈, $\theta \in \Omega = \{1/4, 1/2\}$ 。基於 X , 欲檢定 $H_0 : \theta = 1/2$, vs. $H_a : \theta = 1/4$ 。拒絕域取為 $\{X \leq 3\}$ 。試求此檢定之檢力函數。(解. $K(1/4) = 11/64$, $K(1/2) = 31 \cdot 3^8/4^9$)
17. 某廠牌輪胎之壽命 X 有 $\mathcal{N}(\theta, 5,000^2)$ 分佈。過去經驗顯示 $\theta = 30,000$ 。廠商最近宣稱, 由於製程的改進, θ 應大於30,000, 且很可能 $\theta = 35,000$ 。欲檢定 $H_0 : \theta \leq 30,000$, vs. $H_a : \theta > 30,000$ 。取一組樣本數為 n 之隨機樣本, 並得樣本平均 $\bar{X}$, 拒絕域取為 $\{\bar{X} \geq c\}$ 。試決定 n 之值, 使得檢力函數滿足 $K(30,000) \doteq 0.01$, 且 $K(35,000) \doteq 0.98$ 。(解. $n = 19$ 或 20)
18. 設 $X_1, \dots, X_{12}$ 為一組由 $\mathcal{P}(\theta)$ 分佈所產生之隨機樣本, $\theta \in \Omega = \{\theta | 0 < \theta \leq 1/2\}$ 。欲檢定 $H_0 : \theta = 1/2$, vs. $H_a : \theta < 1/2$ 。拒絕域取為 $\{X_1 + \dots + X_{12} \leq 2\}$ 。試求檢力函數 $K(\theta)$ 在 $\theta = 1/12$ 及 $1/2$ 之值。(解. $K(1/12) \doteq 0.920$, $K(1/2) \doteq 0.062$)
19. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本, σ^2 為已知。欲檢定 $H_0 : \mu = \mu_1$, vs. $H_a : \mu = \mu_2$ 。試證只要 n 夠大, $K(\mu_1)$ 可任意小, 而 $K(\mu_2)$ 可任意大。
20. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{U}(0, \theta)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta \leq \theta_0$, vs. $H_a : \theta > \theta_0$ 。取拒絕域為 $\{X_{(n)} \geq c\}$ 。
(i) 試求檢力函數 $K(\theta)$, 並證明此為 θ 之一增函數;

- (ii) 設 $\theta_0 = 1/2$, 試求 c 之值, 使顯著水準為 0.05;
- (iii) 試決定 n 要多大, 使得對(ii)中之檢定, $K(0.75) = 0.98$;
- (iv) 設 $n = 20$, 且 $X_{(n)} = 0.48$, 試求此時之 p -值。
21. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{E}(\lambda)$ 分佈所產生之隨機樣本。令 $\mu = 1/\lambda$ 。欲檢定 $H_0: \mu \leq \mu_0$, vs. $H_a: \mu > \mu_0$ 。
- (i) 試證對 $\forall 0 < \alpha < 1$, 拒絕域 $\{\bar{X} \geq \mu_0 \chi_{1-\alpha, 2n}^2 / (2n)\}$, 為一顯著水準 α 之檢定;
- (ii) 試以 χ_{2n}^2 之分佈函數 F , 表示此檢定之檢力函數。
- (解. (ii) $1 - F(\mu_0 \chi_{1-\alpha, 2n}^2 / \mu)$)
22. 投擲一公正的銅板 $100n$ 次, 其中 n 為一正整數, 得到投擲數之 52% 個正面。令 p 表銅板出現正面之機率。欲檢定 $H_0: p = 0.5$, vs. $H_a: p \neq 0.5$, 取 $\alpha = 0.01$ 。利用常態近似, 並採連續的更正, 求最小的 n , 使能拒絕 H_0 。(解. $n = 42$)

第8.3節

- 試證(3.1)式中之 Z 有 $\mathcal{N}(0, 1)$ 分佈。
- 試證當 $\sigma_1^2 = \sigma_2^2$, (3.2)式中之 T 有 $\mathcal{T}_{n_1+n_2-2}$ 分佈。
- 試證例3.3中之 p -值約為 0.07。
- 在例3.3中, 分別於(i) $n = 5$ 及(ii) $n = 20$, 於得到 $S = 1.2$ 時, 在 $\alpha = 0.05$ 之下, 對 H_0 做一檢定。並說明此結果是否合理。
- 設 $X_1, \dots, X_{18}$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本。欲檢定 $H_0: \sigma^2 = 0.36$, vs. $H_a: \sigma^2 > 0.36$ 。假設樣本變異數 $S^2 = 0.68$ 。試在 $\alpha = 0.05$ 之下, 作一檢定。(解. $K \doteq 32.1$, $\chi_{0.05, 17}^2 \doteq 27.59$, 拒絕 H_0)
- 設 $X_1, \dots, X_{20}$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本。欲檢定 $H_0: \sigma^2 = 225$, vs. $H_a: \sigma^2 \neq 225$ 。假設樣本變異數 $S^2 = 397$ 。試在 $\alpha = 0.05$ 之下, 作一檢定。(解. $K \doteq 33.5$, 拒絕 H_0)

7. 設 $X_1, \dots, X_{30}$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本, μ 設為未知。欲檢定 $H_0: \sigma^2 = 8$, vs. $H_a: \sigma^2 < 8$ 。假設樣本變異數 $S^2 = 6.4$ 。試在 $\alpha = 0.05$ 之下, 作一檢定。(解. $K = 23$, 不能接受 H_0)
8. 分別對 9 位男學生及 9 位女學生作某項測驗, 男生之成績為 88, 68, 77, 82, 63, 80, 78, 71, 72, 女生之成績為 73, 77, 67, 74, 74, 64, 71, 71, 72。假設男生、女生的成績分別可視為由變異數 σ_1^2, σ_2^2 之常態分佈所產生之隨機樣本, 且兩組樣本假設為獨立。欲檢定 $H_0: \sigma_1^2 = \sigma_2^2$, vs. $H_a: \sigma_1^2 > \sigma_2^2$ 。試在 $\alpha = 0.05$ 之下, 作一檢定。(解. $S_1^2/S_2^2 \doteq 3.86 > 3.44 = f_{0.95, 8, 8}$, 拒絕 H_0)
9. 設 $X_1, \dots, X_9$ 為一組由 $\mathcal{N}(\mu_1, \sigma_1^2)$ 分佈所產生之隨機樣本, $Y_1, \dots, Y_9$ 為一組由 $\mathcal{N}(\mu_2, \sigma_2^2)$ 分佈所產生之隨機樣本。又設觀測到 $\bar{X} = 16$, $\bar{Y} = 10$, $S_1^2 = 36$, $S_2^2 = 45$ 。
- (i) 設 $\sigma_1^2 = \sigma_2^2$, 試給 $\alpha = 0.10$ 下之 $H_0: \mu_1 = \mu_2$, vs. $H_a: \mu_1 \neq \mu_2$ 之檢定;
- (ii) 試給 $\alpha = 0.05$ 下之 $H_0: \sigma_2^2/\sigma_1^2 \leq 1$, vs. $H_a: \sigma_2^2/\sigma_1^2 > 1$ 之檢定。
- (解. (i) $-2 < -1.746$, 拒絕 H_0 , (ii) $0.29 < 0.8$, 接受 H_0)

第 8.4 節

1. 試證例 4.4 中所得之拒絕域, 與例 4.2 中所得相同。
2. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu_1, \sigma^2)$ 分佈所產生之隨機樣本, $Y_1, \dots, Y_m$ 為一組由 $\mathcal{N}(\mu_2, \sigma^2)$ 分佈所產生之隨機樣本, 且兩組隨機樣本假設為獨立。試證
- (i) $(\bar{X} - \bar{Y} - (\mu_1 - \mu_2))/\sqrt{\sigma^2/n + \sigma^2/m}$ 有 $\mathcal{N}(0, 1)$ 分佈;
- (ii) $((n-1)S_1^2 + (m-1)S_2^2)/\sigma^2$ 有 χ_{n+m-2}^2 分佈;
- (iii) $\frac{(nm/(n+m))^{1/2}(\bar{X} - \bar{Y} - (\mu_1 - \mu_2))}{(((n-1)S_1^2 + (m-1)S_2^2)/(n+m-2))^{1/2}}$ 有 $\mathcal{T}_{n+m-2}$ 分佈。

3. 設 $X_{1j}, X_{2j}, \dots, X_{n_j j}$ 為一組由 $\mathcal{N}(\mu_j, \sigma^2)$ 分佈所產生之隨機樣本, $j = 1, \dots, k$, 令 $N = \sum_{j=1}^k n_j$ 。所有 N 個樣本假設皆相互獨立。又設 $\mu_1, \dots, \mu_k$ 及 σ^2 皆未知, 欲檢定 $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$, vs. H_a : 至少有一對 $\mu_r \neq \mu_s, r \neq s$ 。

(i) 試證LRT統計量為

$$\lambda(\mathbf{x}) = \left(\frac{\sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_{\cdot j})^2}{\sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2} \right)^{N/2},$$

其中 $\bar{x}_{\cdot j} = \sum_{i=1}^{n_j} x_{ij} / n_j$, $\bar{x} = \sum_{j=1}^k \sum_{i=1}^{n_j} x_{ij} / N = \sum_{j=1}^k n_j \bar{x}_{\cdot j} / N$;

(ii) 令 $V_1 = \sum_{j=1}^k \sum_{i=1}^{n_j} (X_{ij} - \bar{X})^2 / \sigma^2$, $V_2 = \sum_{j=1}^k \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_{\cdot j})$, $V_3 = \sum_{j=1}^k n_j (\bar{X}_{\cdot j} - \bar{X})^2 / \sigma^2$ 。試證 $V_1 = V_2 + V_3$, V_2 與 V_3 獨立, 又在 H_0 之下, $V_1 \sim \chi_{N-1}^2$, $V_2 \sim \chi_{N-k}^2$, 且 $V_3 \sim \chi_{k-1}^2$;

(iii) 試證LRT檢定之拒絕域等價於 $\{F \geq c\}$, 其中

$$F = \frac{\sum_{j=1}^k n_j (\bar{X}_{\cdot j} - \bar{X})^2 / (k-1)}{\sum_{j=1}^k \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_{\cdot j})^2 / (N-k)},$$

且 F 有 $\mathcal{F}_{k-1, N-k}$ 分佈。又當顯著水準為 α , $C = f_{1-\alpha, k-1, N-k}$ 。

4. 在例4.6中, 取 $n = m = 8$, $\bar{x} = 75.2$, $\bar{y} = 78.6$, $\sum_{i=1}^8 (x_i - \bar{x})^2 = 71.2$, $\sum_{i=1}^8 (y_i - \bar{y})^2 = 54.8$ 。若採用該例中之檢定法, 試決定在 $\alpha = 0.05$ 之下, 是否接受 H_0 。(解. 因 $|T| = 2.27 > 2.145$, 拒絕 H_0)
5. 投擲一出現正面機率為 p 之銅板100次, 得到60個正面。試利用LRT, 及適當的極限結果, 在 $\alpha = 0.1$ 之下, 檢定 $H_0 : p = 1/2$, vs. $H_a : p \neq 1/2$ 。(解. 拒絕 H_0)
6. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{U}[0, \theta]$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta = \theta_0$, vs. $H_a : \theta \neq \theta_0$ 。試給一 α 下之LRT。(解. 拒絕域為 $\{X_{(n)} < \theta_0 \alpha^{1/n}\}$)
7. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{E}(\theta)$ 分佈所產生之隨機樣本。

(i) 試給一 α 下之 $H_0 : \theta = \theta_0$, vs. $H_a : \theta \neq \theta_0$ 的LRT, 又當 n 很大時, 利用定理4.2, 得到近似的拒絕域;

(ii) 試給一 α 下之 $H_0 : \theta = \theta_0$, vs. $H_a : \theta > \theta_0$ 的LRT。

(解. (i) 近似的拒絕域為 $\{-2n(\log(\bar{X}/\theta_0) - \bar{X}/\theta_0 + 1) \geq \chi_{1-\alpha,1}^2\}$,

(ii) 拒絕域為 $\{2n\bar{X}/\theta_0 \geq \chi_{1-\alpha,2n}^2\}$)

8. 設 $X_1, \dots, X_n$ 為一組由 $Be(\theta, 1)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta = \theta_0$, vs. $H_a : \theta \neq \theta_0$ 。試給一 α 下之LRT。又當 n 很大時, 利用定理4.2, 得到近似的拒絕域。

(解. 近似的拒絕域為 $\{-2n(\log(\theta_0/\hat{\theta}) - \theta_0/\hat{\theta} + 1) \geq \chi_{1-\alpha,1}^2\}$, $\hat{\theta} = -n/(\log \prod_{i=1}^n X_i)$)

第8.5節

1. 設 X 有 $B(5, \theta)$ 分佈。欲檢定 $H_0 : \theta = 1/2$, vs. $H_a : \theta = 3/4$ 。試分別求基於 X , (i) $\alpha = 1/32$, (ii) $\alpha = 6/32$, 下之MP檢定。(解. (i) 拒絕域為 $\{X = 5\}$, (ii) 拒絕域為 $\{X = 4, 5\}$)

2. 設隨機變數 X 之p.d.f.為 $f(x|\theta)$, $\theta \in \Omega = \{\theta_0, \theta_1\}$ 。又設 $f(x|\theta)$ 如下表:

$\theta \setminus x$	0	1	2	3
θ_0	0.2	0.3	0.1	0.4
θ_1	0.2	0.25	0.35	0.2

基於 X , 欲檢定 $H_0 : \theta = \theta_0$, vs. $H_a : \theta = \theta_1$ 。

(i) 試列出所有可能的拒絕域, 寫出其對應的 α 與 $K(\theta_1)$ 之值, 並對不同的 α , 給出UMP檢定;

(ii) 試利用奈曼-皮爾生引理以得UMP檢定。

(解. (i) 共有16個可能的拒絕域。UMP檢定如下: $0 \leq \alpha < 0.1$, $C = \emptyset$; $0.1 \leq \alpha < 0.3$, $C = \{X = 2\}$; $0.3 \leq \alpha < 0.4$, $C = \{X = 0, 2\}$; $0.4 \leq \alpha < 0.6$, $C = \{X = 1, 2\}$; $0.6 \leq \alpha < 0.8$, $C = \{X = 0, 1, 2\}$;

$0.8 \leq \alpha < 1$, $C = \{X = 0, 1, 2\}$ 或 $\{X = 1, 2, 3\}$; $\alpha = 1$, $C = \{X = 0, 1, 2, 3\}$; (ii) $k > 3.5$, $C = \emptyset$, $\alpha = 0$; $1 < k \leq 3.5$, $C = \{X = 2\}$, $\alpha = 0.1$; $5/6 < k \leq 1$, $C = \{X = 0, 2\}$, $\alpha = 0.3$; $1/2 < k \leq 5/6$, $C = \{X = 0, 1, 2\}$, $\alpha = 0.6$; $0 \leq k \leq 1/2$, $C = \{X = 0, 1, 2, 3\}$, $\alpha = 1$)

3. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\theta, 1)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta = 0$, vs. $H_a : \theta = 1$ 。

(i) 試求一MP檢定;

(ii) 取 $\alpha = 0.05$, $n = 25$, 試求此時之MP檢定, 及檢力。

(解. (i) 拒絕域為 $\{\bar{X} > c\}$, (ii) $c = 0.329$, 檢力為 0.9996)

4. 重做上題, 改為檢定 $H_0 : \theta = 0$, vs. $H_a : \theta = -1$ 。(解. (i) 拒絕域為 $\{\bar{X} < c\}$, (ii) $c = -0.329$, 檢力為 0.9996)

5. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x)$, $x = 0, 1, \dots$, 所產生之隨機樣本。欲檢定 $H_0 : f(x) = e^{-1}/x!, x = 0, 1, \dots$, vs. $H_a : f(x) = (1/2)^{x+1}, x = 0, 1, \dots$ 。

(i) 試證取拒絕域 $\{\log 2 \sum_{i=1}^n X_i - \log(\prod_{i=1}^n X_i!) \leq c\}$ 為一MP檢定;

(ii) 在(i)中, 設 $n = 1$, 且 $c = \log(4/3)$ 。試證此時拒絕域為 $\{X_1 = 0, 3, 4, 5, \dots\}$, 並求此時型 I 及型 II 錯誤之機率 α, β 。

(解. (ii) α 約 0.448, $\beta = 0.375$)

6. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(0, \sigma^2)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \sigma^2 = 1$, vs. $H_a : \sigma^2 = 2$ 。

(i) 試求一 $\alpha = 0.05$ 下之MP檢定;

(ii) 試問若欲檢定 $H_0 : \sigma^2 = 1$, vs. $H_a : \sigma^2 = 4$, 則(i)中之拒絕域, 是否為 $\alpha = 0.05$ 下之最佳拒絕域?

(iii) 試問若欲檢定 $H_0 : \sigma^2 = 1$, vs. $H_a : \sigma^2 = \sigma_1^2 > 1$, 則(i)中之拒絕域是否為 $\alpha = 0.05$ 下之最佳拒絕域?

(解. (i) 拒絕域為 $\{\sum_{i=1}^n X_i^2 > 18.3\}$, (ii) 是, (iii) 是)

7. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\theta_1, \theta_2)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta_1 = 0, \theta_2 = 1$, vs. $H_a : \theta_1 = 1, \theta_2 = 4$ 。試求一最佳檢定。(解. 拒絕域為 $\{3 \sum_{i=1}^n X_i^2 + 2 \sum_{i=1}^n X_i > c\}$)
8. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\theta, 100)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta = 75$, vs. $H_a : \theta = 78$ 。試證拒絕域 $\{\bar{X} > c\}$ 為一最佳檢定。並求 n 及 c , 使得型 I 錯誤及型 II 錯誤之機率, 分別約為 0.05 及 0.10。(解. $n = 95$ 或 96, $c = 76.7$)
9. 設 $X_1, \dots, X_n$ 為一組由 $\text{Ber}(p)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : p = 1/2$, vs. $H_a : p = 1/3$ 。試證拒絕域 $\{\sum_{i=1}^n X_i \leq c\}$ 為一最佳檢定。並利用中央極限定理求 n 及 c , 使得型 I 錯誤及型 II 錯誤之機率, 分別約為 0.10 及 0.20。(解. $n = 38$ 或 39, $c = 15$)
10. 設 $X_1, \dots, X_{10}$ 為一組由 $\mathcal{P}(\theta)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta = 0.1$, vs. $H_a : \theta = 0.5$ 。試證 $\{\sum_{i=1}^{10} X_i \geq k\}$ 為一最佳拒絕域。並求 $k = 3$ 時之型 I 及型 II 錯誤之機率 α, β 。(解. $\alpha = 0.08, \beta = 0.125$)
11. 設 $X_1, \dots, X_{10}$ 為一組由 $\text{Ber}(\theta)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta = 1/4$, vs. $H_a : \theta < 1/4$ 。取拒絕域為 $\{\sum_{i=1}^{10} X_i \leq 1\}$ 。試求檢力函數 $K(\theta)$, $0 < \theta \leq 1/4$ 。(解. $K(\theta) = (1 - \theta)^9(1 + 9\theta)$, $0 < \theta \leq 1/4$)
12. 設 X_1, X_2, X_3, X_4 為一組由 $\mathcal{U}(0, \theta)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta = 1$, vs. $H_a : \theta \neq 1$ 。令 $X_{(4)} = \max\{X_1, X_2, X_3, X_4\}$ 。取拒絕域為 $\{X_{(4)} < 1/2 \text{ 或 } X_{(4)} > 1\}$ 。試求檢力函數 $K(\theta)$, $\theta > 0$ 。(解. $K(\theta) = 1, 0 < \theta \leq 1/2; 1/(16\theta^4), 1/2 < \theta < 1; 1 - 15/(16\theta^4), \theta \geq 1$)
13. 設 $X_1, \dots, X_n$ 及 $Y_1, \dots, Y_n$ 分別為由 $\mathcal{N}(\mu_1, 400)$ 及 $\mathcal{N}(\mu_2, 225)$ 分佈所產生之隨機樣本。令 $\theta = \mu_1 - \mu_2$ 。欲檢定 $H_0 : \theta = 0$, vs. $H_a : \theta > 0$ 。令 $\bar{X} = \sum_{i=1}^n X_i/n, \bar{Y} = \sum_{i=1}^n Y_i/n$ 。取拒絕域為 $\{\bar{X} - \bar{Y} > c\}$, 令 $K(\theta)$ 表檢力函數。試決定 n 及 c 之值, 使得 $K(0)$ 及 $K(10)$ 分別約為 0.05 及 0.90。(解. $n = 53$ 或 54, $c = 5.6$)

14. 設 $X_1, \dots, X_n$ 為一組由 $Be(\theta, 1)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta = 1$, vs. $H_a : \theta = 2$ 。
- (i) 設 $n = 1$, 試給一 $\alpha = 0.05$ 下之MP檢定;
- (ii) 試求(i)中 $\theta = 2$ 時之檢力;
- (iii) 對一般的 n , 試給一 α 下之MP檢定。
- (解. (i) 拒絕域為 $\{X > 0.95\}$, (ii) 0.0975, (iii) 拒絕域為 $\{\prod_{i=1}^n X_i > c\}$, 其中 $P(\prod_{i=1}^n X_i > c | \theta = 1) = 0.05$)
15. 設 X 有 $\mathcal{NB}(r, 1/2)$ 分佈。基於 X , 欲檢定 $H_0 : r = 1$, vs. $H_a : r = 2$ 。試給一 $\alpha = 0.125$ 下之MP檢定, 並求 $r = 2$ 時之檢力。(解. 拒絕域為 $\{X \geq 4\}$, 檢力 = 0.5)
16. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\theta, \sigma^2)$ 分佈所產生之隨機樣本, 其中 σ^2 為已知。欲檢定 $H_0 : \theta = 0$, vs. $H_a : \theta = 1$ 。
- (i) 試證給定型 I 及型 II 錯誤之機率 α, β , 能決定 n 之值;
- (ii) 利用(i), 當 $\alpha = 0.05, \beta = 0.1, \sigma = 1$, 試求 n 之值。
- (解. (ii) $n = 9$)
17. 設 $X_1, \dots, X_{100}$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \mu = 3, \sigma^2 = 4$, vs. $H_a : \mu = 3.5, \sigma^2 = 4$ 。試在 $\alpha = 0.01$ 之下給一MP檢定。又問當 $\bar{X} = 3.2$ 時, 是否接受 H_0 。(解. 拒絕域為 $\{\bar{X} > 3.466\}$, 接受 H_0)
18. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(0, \sigma^2)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \sigma = \sigma_0$, vs. $H_a : \sigma = \sigma_1$, 其中 $\sigma_1 < \sigma_0$ 。試證拒絕域 $\{\sum_{i=1}^n X_i^2 > c\}$ 為一MP檢定。
19. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{P}(\lambda)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \lambda \leq \lambda_0$, vs. $H_a : \lambda > \lambda_0$ 。試證拒絕域 $\{\sum_{i=1}^n X_i > k\}$, 其中 k 由 $P(T > k | \lambda = \lambda_0) = \alpha$ 所決定, 為一 α 下之UMP檢定。
20. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x|\theta) = (\theta + 1)^{-1}(\theta/(\theta + 1))^x, x = 0, 1, \dots$, 所產生之隨機樣本, $\theta > 0$ 。欲檢定 $H_0 : \theta = \theta_0$, vs. $H_a :$

$\theta > \theta_0$ 。試求 $-\alpha$ 下之UMP檢定。(解.拒絕域為 $\{\sum_{i=1}^n X_i \geq c\}$, 其中 $P(\sum_{i=1}^n X_i \geq c | \theta = \theta_0) = \alpha$)

21. 試以系理5.2重做例5.7。

22. 設 X 有 $\mathcal{B}(n, \theta)$ 分佈, $0 < \theta < 1$ 。試證 $\frac{d}{d\theta}P(X \geq k) = n \binom{n-1}{k-1} \theta^{k-1} (1-\theta)^{n-k}$, $1 \leq k \leq n$ 。由此得 $P(X \geq k)$ 為 $-\theta$ 之漸增函數。

23. 試證底下各分佈族皆有MLR。

(i) $\mathcal{N}(\theta, \sigma^2)$ 分佈族, σ^2 已知;

(ii) $P(\theta)$ 分佈族;

(iii) $\mathcal{B}(n, \theta)$ 分佈族, n 已知。

24. 設 X 有 $\mathcal{C}(0, \theta)$ 分佈, $\theta > 0$ 。

(i) 試證此分佈族並無MLR;

(ii) 試證 $|X|$ 為 θ 之一充分統計量, 且 $|X|$ 之分佈有MLR。

25. 設 $f(x|\theta)$ 為羅吉斯分佈之p.d.f., 且

$$f(x|\theta) = \frac{e^{(x-\theta)}}{(1 + e^{(x-\theta)})^2}, -\infty < x < \infty, -\infty < \theta < \infty。$$

(i) 試證分佈族 $\{f(x|\theta), \theta \in R\}$ 有MLR;

(ii) 基於一個觀測值 X , 欲檢定 $H_0 : \theta = 0$, vs. $H_a : \theta = 1$, 試求 $-\alpha$ 下之MP檢定;

(iii) 欲檢定 $H_0 : \theta \leq 0$, vs. $H_a : \theta > 0$, 試證(ii)中所得之檢定, 為 $-\alpha$ 下之UMP檢定。

26. 設 X 有 $\mathcal{C}(\theta, 1)$ 分佈, $\theta \in R$ 。

(i) 試證此分佈族並無MLR;

(ii) 欲檢定 $H_0 : \theta = 0$, vs. $H_a : \theta = 1$, 試證拒絕域 $\{1 < X < 3\}$ 為 $-\alpha$ 下之MP檢定;

(iii) 試證或否證: (ii)中所得之檢定, 並非 $H_0 : \theta \leq 0$, vs. $H_a : \theta > 0$ 之UMP檢定。

27. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\theta, \sigma^2)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta \leq \theta_0$, vs. $H_a : \theta > \theta_0$ 。

(i) 若 σ^2 已知, 試證拒絕域 $\{\bar{X} > \theta_0 + z_{1-\alpha}\sigma/\sqrt{n}\}$ 為一顯著水準 α 之檢定, 並證明此檢定可由 LRT 導出;

(ii) 試證(i)中之檢定為 UMP 檢定;

(iii) 若 σ^2 未知, 試證拒絕域 $\{\bar{X} > \theta_0 + t_{1-\alpha, n-1}S/\sqrt{n}\}$ 為一顯著水準 α 之檢定。並證明此檢定可由 LRT 所導出。

28. 設 $X_1, \dots, X_n$ 為一組隨機樣本。試證在下述各情況中, 皆有一隨機變數 $V = V(X_1, \dots, X_n)$ 之分佈族有 MLR。假設共同分佈為

(i) $\Gamma(\alpha, \theta)$, α 已知, $\theta \in \Omega = (0, \infty)$;

(ii) $\Gamma(\alpha, 1/\theta)$, α 已知, $\theta \in \Omega = (0, \infty)$;

(iii) $\mathcal{NB}(r, \theta)$, r 已知, $\theta \in \Omega = (0, 1)$ 。

(解. (i) $V = \sum_{i=1}^n X_i$, (ii) $V = -\sum_{i=1}^n X_i$, (iii) $V = -\sum_{i=1}^n X_i$)

29. 某地的年降雨量可假設有常態分佈, 標準差 $\sigma = 3$, 而期望值 μ 未知。過去十年的年降雨記錄為 30.5, 34.1, 27.9, 29.4, 35.0, 26.9, 30.2, 28.3, 31.7, 25.8 (單位為公分)。欲檢定 $H_0 : \mu = 30$, vs. $H_a : \mu < 30$ 。試在 $\alpha = 0.05$ 之下, 給一 UMP 檢定。(解. 拒絕域為 $\{\bar{X} < 28.44\}$, 接受 H_0)

30. 自某大學中隨機地抽取 400 位學生, 其中有 95 位學過鋼琴。令 θ 表大學生中學過鋼琴所佔的比例。試在 $\alpha = 0.05$ 之下, 檢定 $H_0 : \theta = 0.25$, vs. $H_a : \theta < 0.25$ 。(解. 拒絕域為 $\{X < 86\}$, 接受 H_0)

31. 設 $X_1, \dots, X_n$ 為一組由 $\text{Ber}(p)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : p \leq 1/2$, vs. $H_a : p > 1/2$, 且要求 $\alpha = 0.05$, $K(7/8) = 0.95$ 。試藉助中央極限定理, 決定所需之樣本數。(解. $n = 14$)

32. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{E}(\theta)$ 分佈所產生之隨機樣本。欲檢定 $H_0 : \theta \leq \theta_0$, vs. $H_a : \theta > \theta_0$ 。

- (i) 試在 α 之下, 給一UMP檢定;
- (ii) 設 $\alpha = 0.05$, $\theta_0 = 0.001$ 。試決定 n 之值, 使得 $K(0.002) \geq 0.95$ 。
- (解. (i) 拒絕域為 $\{\sum_{i=1}^n X_i < c\}$, $P_{\theta_0}(\sum_{i=1}^n X_i < c) = \alpha$; (ii) $n = 23$)
33. 設 $X_1, \dots, X_n$ 為由 $\mathcal{N}(\theta, \sigma^2)$ 分佈所產生之隨機樣本, σ^2 為未知。欲檢定 $H_0 : \theta = \theta_0$, vs. $H_a : \theta \neq \theta_0$ 。試證拒絕域 $\{|\bar{X} - \theta_0| > t_{1-\alpha/2, n-1} S / \sqrt{n}\}$ 為一顯著水準 α 之UMP。
34. 試證定理5.3中之 $K(\theta)$ 為 θ 之非漸增函數。

參考文獻

1. 黃文璋(1999a). 數學欣賞。華泰文化事業股份有限公司, 台北。
2. 黃文璋(1999b). 純屬巧合。數學傳播季刊, 第23卷第4期, 6-21。
3. 葉偉文譯(2001). 統計, 改變了世界(Salsburg, D.原著: The Lady Tasting Tea)。天下遠見出版股份有限公司, 台北。
4. 羅夢娜(1987a). 實驗設計淺談。科學月刊, 第18卷第5期, 334-337。
5. 羅夢娜(1987b). 百農嘗一草—實驗設計之應用。科學月刊, 第18卷第5期, 338-340。
6. Fisher, R.A.(1935). *The Design of Experiments*. Oliver and Boyd, Edinburgh, United Kingdom.
7. Neyman, J.(1950). *A First Course in Probability and Statistics*. Henry Holt & Co., New York.