統計導論

〔美〕赫尔 波特 等原著

曉園出版社
曉園出版公司

統計導論

原著者 Hoel・Port・Stone
譯著者 詹 世 煌

曉 園 出 版 社
世界圖書出版公司
北京・广州・上海・西安
1992

内 容 简 介

本书是为大学一学期的数理统计课程而设计的。本书从理论的观点，以一种基本的、有系统的方式叙述数理统计学。本书还附有习题解答，便于自学。

统计导论

赫尔 波特等原著

詹世煌 译著

*

世界图书出版公司北京重印

北京朝阳门内大街 137 号

通州印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1992 年 10 月第一版 开本：711×1245 1/24

1992 年 10 月第一次印刷 印张：10.5

印数：0001—1400

ISBN：7—5062—1320—6/Z·39

定价：8.20 元（W_B9201/19）

序　言

本書乃是為一學期數理統計課程而設計，其編寫的方式，與我們一系列三本書中的第一册——機率導論（ Introduction to Probability Theory ）有相當密切的關係。它乃假定學者對以基本微積分為預備知識的一學期機率課程中所包含的內容，有相當的認知。

本書的目的，乃是自理論的觀點以一種基本的、有系統的方式，來敘述數理統計學。我們試圖只考慮一些重要的基本概念，且詳細的加以討論，以使學生能對理論的數學及動機有所認知；學過統計學的學生常常對此一主題的重要概念和方法僅有一模糊的見解，我們希望此書對統計方法一致且邏輯的架構，能對學者的觀念有所啓發。

本書理論的鋪陳係依抉擇理論的一些基本概念而來，因此，除了傳統的統計方法外，亦加入了貝氏方法；但由於篇幅所限，我們僅能介紹此等方法的最基本概念。貝氏方法列在每章的最後，因此若時間上不允許，可省略。

統計學上最重要的定理且最有用的方法，乃是關於一般綫性假設的檢定。此一定理為衆多特殊檢定的基礎，且可應用於很多重要問題中。有關此定理之證明很少列於初等水準的教科書中；但是由於此定理的重要性，和修過初等微積分的學生如今正修讀矩陣代數，因此本書以簡單代數和幾何的方法證明了此一定理；此一定理之證明，列於第五章，無疑為本書中最艱深的部分，但其功用却值得我們如此處理。對那些不具備此一代數背景的學生，最好略過此章的證明，而直接討論其應用問題。

雖然基本微積分對機率導論（ Introduction to Probability The-

ory ）言便已足夠，本書之學者仍得有某些基本的矩陣代數的知識。這些知識在第四章和第五章將用到，因為該章中最小平方理論，係以矩陣符號和技巧來敘述。但僅在第五章，方用到比矩陣代數之最簡單概念要深的理念，有關此二章中所需用到的矩陣方法，列於本書附錄中。

某些教師或許會驚訝的發現本書並沒有介紹充分性的概念。充分性在敘述高等統計理論上非常有用，但在本書之初等水準上並沒有什麼用處；本書亦不包含在基本教科書中常見的其他主題，我們之所以省略它們，乃認為若在課堂上討論此等課題，則對本書中所涵蓋的基本內容，便無法以充份的時間詳細討論。

每章之末的習題係依該章中該等素材出現的順序來安排，屬於計算的問題先列出，其答案置於附錄中。

雖然本書係為一週有三次之一學期課程而設計，我們亦可將其內容適當地安排、調整，以適合時間較短課程之用；此只要將無母數方法一章或貝氏方法，或某些證明，或上述之某些組合予以省略，即可達成。標有星號的各節亦可以刪略。若課程較本書所設計者為長則可包含所有的理論部份，並花費較多的時間於習題上。

作者很感激 Frank Samaniego 先生做出習題中的很多解答，同時亦感謝 Gerry Formanack 卓越的打字技巧。

目　錄

第一章

基本原理

在各種科學中，有很多現象係由具機遇性質之法則或關係所支配，此意謂我們可用一機率模式，來適當地描述此現象的發生情形，此等模式我們已於第一册之機率導論（Introduction to Probability Theory）中介紹，並應用於機遇遊戲和其他物理實驗中。雖然物理學所得到的實驗出象似是比非物理學所得者（因此機率模式將更適合）來得穩定，此等模式却與其他學科同樣地適用，其基本差異爲其他學科可能有較多的不可控制變數干擾到所欲研究的變數，因而導致較大的變異性。機率模式應用於各種科學現象上的研究，導致現在我們一般所謂之統計方法的發展。

統計方法所欲解決的簡單典型問題如下：

就某藥品的一批量中取出一些樣本，而後以檢定爲基準決定此批量之品質是否滿意，

以一小批投票者的意見爲基礎來預測選民對某重大問題的偏好。

以一高中學生的成績，和已進入大學之學生的成績爲基準，求該高中生在大學裏表現良好的機會。

分析聲納的訊號，以決定一潛艇是否正接近中。

對此類問題，可由視用來做決策之資料爲一隨機變數X的觀察值，而將它以數學型式列出。X的分配，可視爲當母數值θ特定時，屬於某分配族的一分配。此問題爲在此資料爲基礎下，決定該密度族中，何者可表X的分配，此稱爲統計推論問題（the problem of statistical inference）。例如，在以投票法決定選民對某一問題偏好的問題中，變數X可視爲一不連續隨機變數，其值依該選民之贊成或不贊成此一問題而爲1或0，而母數θ表示選民中贊成此一問題之未知比率。統計推論之典型問題爲，在民意測驗專家所得的一組觀察值$(x_1,\cdots,x_n)$爲基礎下，決定θ值是否大於.60

統計方法在範圍上遠比此地所將敍述的統計推論問題來得廣，它包括

如何設定實驗，用何種模式較適當和如何用情報諸問題。但是，求取對隨機數之分配做推論的最佳，我們將幾乎完全排除。此意是說，我們設隨機變數X已具有依存於未知母數θ之某一分配形態，而我們的目的，乃是對此θ做某種推論。在大部份問題中，θ將爲機率密度函數$f(x|\theta)$之顯然母數；但是，θ亦可能只是一個指標，以區別此類密度函數族中之各元素。隨機變數X和母數θ，可爲具多個成份的向量變數，但在我們基本原理的討論中，爲簡化敍述起見，皆將其視爲一維來處理，至若推廣到向量變數的情形，則留待下一章再考慮。有關此類典型機率模式，爲以θ表一實驗之單一試行中成功機率的二項分配，θ表分配之均數的卜松分配，和具已知變異數，且θ表分配之均數的常態分配；若一常態分配之均數和變異數皆未知，則θ將表母數向量(μ,σ)。

在形如θ爲選民偏好某一問題之比率的統計推論問題上，於抽取選民之時，母數θ視爲固定但未知之常數。但在某些問題中，母數θ可視爲具有一已知機率分配之隨機變數。若如此，我們設此一已知的分配有密度函數$\pi(\theta)$，函數$f(x|\theta)$因而表變數θ固定下的條件分配，而X和θ的聯合密度爲$f(x,\theta)=\pi(\theta)f(x|\theta)$，此地之密度乃用來表不連續和連續隨機變數，此種表示法已於第一册中使用，本書中亦將繼續使用。由於通常大寫字母通常表一隨機變數，而其對應之小寫字母表它的值，但在本書中同時用θ表示隨機變數和它的值，因此在符號上有些微的不一致。當θ視爲隨機變數時，它的機率分配稱爲事前分配 (prior distribution)，之所以如此稱呼，乃因在實驗而對θ做推論之前，$\pi(\theta)$已知之故。爲說明θ可視爲一隨機變數的問題，設我們欲檢定來自一批量的樣本，以決定該批量中的瑕疵品比率θ是否超過某容受比率θ_0。

假設此購買廠商定期地收到此等批量，且由過去批量之θ'，已決定了θ的分配，則此批量的θ可視爲具有該等分配之隨機變數的值，只是此θ值現在未知。在統計理論之大部份古典方法中，皆視θ爲一常數，且僅依存於隨機變數X的一組觀察值來做推論。此部份源於在社會和生命科學中（引出大部份理論之處），沒有$\pi(\theta)$可用或適用。當然，若有θ之事前情報可用而予以忽視，將會造成錯誤，即使該等情報不能表成精確的機率分配亦如此。

要言之，我們將設已有依存於母數 θ 之隨機變數 X 的機率分配，且我們希望在隨機變數 X 的一些觀察值和 θ 的事前分配（若有的話）爲基礎下，對 θ 做推論。

1－1　問題的形式

由於推論之做成乃由隨機變數 X 的一組觀察值 $x_1, \cdots x_n$ 而來，我們必須引介這些值的函數 $d = d(x_1, \cdots x_n)$ 以做推論，此一函數稱爲決策函數 (decision function)。此一函數之性質依對 θ 所欲做之推論種類而定，其最簡單之例爲我們只想知道某敍述爲眞或爲僞。例如，我們欲知一藥品的批量是否超過品質標準，一雷達掃描是否已發現一飛彈，或一學校區域內營養不良的學童數是否超過 10 %。我們將取決策分析中正的觀點，而對每一決策賦予一個行動。故前述問題中，有二個可能行動 a_1 和 a_2，a_1 對應於接受敍述爲眞的抉擇，而 a_2 對應於棄却該抉擇。對一組有 n 個觀察值或樣本值的樣本，我們有 n 維樣本空間。由於一決策函數 $d(x_1, \cdots x_n)$ 必須對此樣本空間中的每一點，決定取行動 a_1 或 a_2，因此此一函數必須恰好分割此樣本空間爲二部分，一部分包含將取 a_1 的樣本點，而另一部份包含將取行動 a_2 的樣本點。例如，若小的 X 值對應於敍述爲眞，則 n 維樣本空間的一可能分法爲，將球體 $x_1^2 + \cdots + x_n^2 = r^2$，此地 r 爲適當選取之常數，內的所有點賦予 a_1 而其他點賦予 a_2。前述只有二種可能行動之類的問題稱爲假設檢定問題 (hypothesis testing problems)。

若可能行動超過二個，則決策問題將較複雜。例如，設已知歐洲之某區域曾居住五個不同的種族，一人類學家欲就該地區所發現之一堆骨骸的特性，以決定其屬於那一種族即是。由於假設檢定程序的簡化，含有二個以上可能行動的問題，有時被誤爲假設檢定問題來處理。例如，引進一新藥治療一疾病，我們欲知此新藥在治療此疾病的效果上是優於，劣於或等於舊藥；此時若視此爲一兩個行動問題則不正確；在二個行動問題上，我們只能檢定此新藥在治療該病症上，是否較舊藥有較佳或相等的效果。若一問題具有 $k > 2$ 個可能行動（k 爲有限數），則稱爲多元決策問題 (multiple-decision problems)，此等問題的決策函數必須分割此樣本空間爲 k 部份

，包含第 i 部份的樣本點則賦予其行動 $a_i, i = 1, \cdots, k$。

第三類問題爲，欲對有無限多種可能的母數 θ，預測其值爲何。故若 θ 表選民中將投候選人 A 的比率，則我們得對該比率有一精確的推定值，而非只是決定其是否超過 1/2。此時決策函數 $d(x_1, \cdots\cdots, x_n)$ 爲一實值函數，它的值域理論上可取爲區間〔0,1〕。此類的問題稱爲推定問題（estimation problems）。

爲說明此三類問題如何在同一實驗狀況下產生，設欲比較能降低血壓之兩種新藥的效果。令 X 表病人處理後之血壓與處理前之血壓的比，θ 表關於某類病人之該等隨機變數的平均值。假設檢定之一典型問題爲：以該兩種藥之實驗爲基礎，決定是 $\theta_1 \leq \theta_2$ 或 $\theta_1 > \theta_2$，在此 θ_1 和 θ_2 對應於此二藥。自實務的觀點來看，除非有特殊的利益，我們很少以某藥較另一藥爲佳來比較，通常對此類問題，我們以多元決策問題來考慮，即有如下之三種可能：$\theta_1 - \theta_2 \leq -\delta$，$-\delta < \theta_1 - \theta_2 < \delta$，$\theta_1 - \theta_2 \geq \delta$，在此 δ 爲實務上考慮此檢定有用之最小差異。若由實驗得到（不管取那一種形式）第一種藥較優的抉擇，則得針對此藥做其他的實驗，以得 θ_1 的精確推定值。

1-2 風險函數

欲一決策函數成功的達成它的目的，必須以數值形式來測度它。若一實驗者能夠對各種錯誤決策的嚴重性賦予其權數，則這些權數可定義爲一損失函數 $\mathscr{L}(\theta, a)$（loss function $\mathscr{L}(\theta, a)$）。我們設定此一函數爲當 θ 爲母數眞值時，取行動 a 所導致的懲罰，此懲罰之大小以數值之大小來衡量。例如，若已有具未知均數 θ 之常態變數的觀察值 $x_1, \cdots, x_n$，且我們欲用它們來推定 θ，則行動 a 中推估 θ 之值爲 $d(x_1, \cdots, x_n)$，此決策中誤差的大小爲 $|\theta - d(x_1, \cdots x_n)|$，因而典型之損失函數可爲 $\mathscr{L}(\theta, a) = |\theta - a|$。爲說明損失函數與決策函數和觀察值間的依存關係，我們以 $\mathscr{L}(\theta, d(x_1, \cdots, x_n))$ 來表示。損失與此函數連在一起，乃說明我們的目的在使 $\mathscr{L}$ 爲最小。由於我們希望決策誤差爲最小，顯然地，我們希望 $\mathscr{L}$ 爲誤差減小時，其值亦減小的函數。如何取前述所討論之三類問題的 $\mathscr{L}$，將在 1.4 節加以考慮。

若所能採取的決策在本質上爲屬性時，便產生問題了，因爲在此情況下我們很難對不正確的決策賦予一數值。一人爲重修他的房子，他可能面臨五種色調的選擇，若在此種選擇中，美學之考慮與金錢同等重要，則欲賦予一損失函數將當相不方便。但是，若一人能對每一種可能選擇應用一數值函數，即效用函數，而後就此效用函數賦予一偏好順序，則可解決此類問題。即使是屬量性質的問題，亦常需引進一效用函數，而以數量形式表達某人在各種可能下的偏好。此後我們假設，若一問題必要或需要引進一效用函數，則$\mathscr{L}$即表此一函數。

到現在爲止我們所討論者，乃在有某隨機變數X的一組觀察值$x_1, \cdots x_n$下，試圖取這些值的一良好函數。在評估任意決策函數的有效性時，我們必須盱衡整個的績效，而非只是看對某單一實驗下它多好。現考慮一實驗，且從此實驗之隨機變數X中取得n個觀察值，這些觀察值的可能值以$X_1, \cdots, X_n$表示。若觀察值係隨機抽取，則由第一章中隨機抽樣的定義知，此n個隨機變數獨立，且皆服從與X相同的分配。因此決策函數$d = d(X_1, \cdots, X_n)$爲一隨機變數，而損失函數$\mathscr{L} = \mathscr{L}(\theta, d(X_1, \cdots, X_n))$亦爲一隨機變數。爲衡量一決策函數的總體有效性，我們求此損失函數的期望值，而後用它做爲評估的依據，由此定義了一新函數，稱爲風險函數，如下

定義1 風險函數$\mathscr{R}$的公式爲

$$\mathscr{R}(\theta, d) = E_\theta \mathscr{L}(\theta, d(X_1, \ldots, X_n)),$$

此地乃是在θ固定下，對隨機變數$X_1, \cdots, X_n$的分配取期望值

在隨機抽樣且X具有密度$f(x|\theta)$下，此n個隨機變數的分配爲$\prod_{i=1}^{n} f(x_i|\theta)$，若$X$爲一連續隨機變數，則其風險函數爲

$$(1) \qquad \mathscr{R}(\theta, d) = \int \cdots \int \mathscr{L}(\theta, d(x_1, \ldots, x_n)) \prod_{i=1}^{n} f(x_i \mid \theta)\, dx_1 \cdots dx_n.$$

至若不連續隨機變數，上述之積分可以對所有$x's$之可能值取加總來代替。由於寫出重積分不方便，我們將以一簡寫符號來代替：即若樣本大小爲1，X表基本之隨機變數，而若取樣本大小$n > 1$，則X表隨機變數向量$X = (X_1, \cdots, X_n)$。故(1)中之重積分符號可以較簡潔的方式表爲

$$\mathscr{R}(\theta, d) = \int \mathscr{L}(\theta, d(x)) f(x \mid \theta)\, dx \tag{2}$$

此地之 $f(x|\theta)$ 表向量變數 X 的密度，而 dx 表 $dx_1 \cdots dx_n$。

現在設我們欲以二決策函數 d_1 和 d_2 的風險函數 $\mathscr{R}(\theta, d_1)$ 和 $\mathscr{R}(\theta, d_2)$ 來比較 d_1 和 d_2，此以其圖形來比較將較容易了解。考慮如圖 1 和 2 中的兩組圖形，此二圖形爲該二風險函數之兩種可能情形。在圖 1 中，顯然決策函數 d_1 優於 d_2，因爲對每一 θ 值，它的風險函數值小於 d_2 的風險函數值，而我們的目的乃是使此風險函數爲最小。但在圖 2 中，却沒有一個函數優於另一函數，此乃因對某些 θ 值，函數 d_2 優於 d_1，而對其他值，却 d_1 優於 d_2 之故。

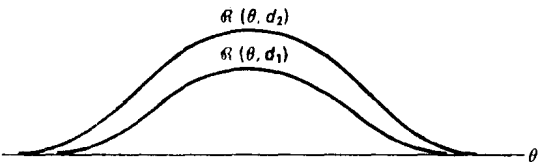

圖 1

很不幸的，圖 -1 中所示的情形很少發生，至少欲對所取的決策函數間以情報做比較時如此。欲就如圖2中所示的狀況（此爲一較典型的狀況），而在 d_1 和 d_2 間做一選擇是困難的。在此情形下，我們必須另外引進一些原則或評判標準，以便有所取捨。在此等原則中，較爲人所熟知者爲以風險函數之最大值爲比較的一個評判標準。若某一風險函數比另一個有較小的最大值，則將此較小最大值的決策函數視爲較佳。若有一決策函數 d 的風險函數，其最大值在所有與之比較的風險函數中有最小的值，則 d 稱爲在大中取小意味下的最佳決策函數，故

定義2 若函數 d_0 滿足

$$\max_{\theta} \mathscr{R}(\theta, d_0) = \min_{d \in D} \max_{\theta} \mathscr{R}(\theta, d).$$

則稱 d_0 為決策函數群中的大中取小決策函數(minimax decision function)

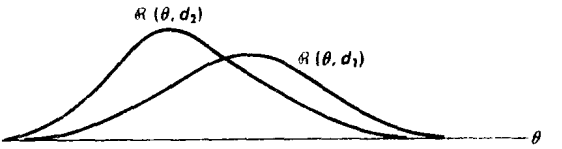

圖 2

由圖 -2 可知，d_1 在大中取小意味下，比之 d_2 為一較佳決策函數，因其風險函數的最大值比之 d_2 的風險函數小故也。若決策函數群 D 中僅含此二函數，則 d_1 為關於此群之大中取小決策函數。

引介其他原則之優點，如大中取小，乃在於它將決策函數的比較導向實數之比較。我們尚有其他原則可以介紹，惟此地留待需要它們時再討論。

為說明前述概念，考慮以單一觀察值 X 為基礎，來推定卜松分配之均數 θ 的問題。現在 $f(x|\theta) = (e^{-\theta}\theta^x)/x!, x = 0, 1, \cdots$，我們取決策函數為 $d(x) = cx$，c 為某正常數，且設損失函數為 $\mathscr{L}(\theta, d) = (d-\theta)^2/\theta$。依公式(1)風險函數為下列之和

$$\mathscr{R}(\theta, d) = \sum_{x=0}^{\infty} \frac{(cx - \theta)^2}{\theta} \frac{e^{-\theta}\theta^x}{x!}.$$

$\mathscr{R}(\theta, d)$ 的計算不難，但若視 $\mathscr{L}$ 的期望值為 $\mathscr{R}$ 的基本定義，且應用第一冊中所導出之 E 的性質，我們可求 $E(cX - \theta)^2/\theta$.，而此期望值更易求得，蓋若將 $(cX-\theta)^2/\theta$ 寫成如下之形式

$$\frac{(cX - \theta)^2}{\theta} = \frac{c^2}{\theta}\left(X - \frac{\theta}{c}\right)^2 = \frac{c^2}{\theta}\left(X - \theta + \theta\left(1 - \frac{1}{c}\right)\right)^2$$

$$= \frac{c^2}{\theta}\left\{(X - \theta)^2 + 2\theta\left(1 - \frac{1}{c}\right)(X - \theta) + \theta^2\left(1 - \frac{1}{c}\right)^2\right\}.$$

則因 X 為一卜松變數，而由第一冊知 X 的均數和變異數皆為 θ，因而 $E(X - \theta) = 0$ 且 $E(X-\theta)^2 = \theta$。應用此一事實於前述之和，可得

$$(3) \qquad \mathcal{R}(\theta, d) = \frac{c^2}{\theta}\left\{\theta + \theta^2\left(1 - \frac{1}{c}\right)^2\right\} = c^2 + \theta(c-1)^2.$$

現在考慮 c 應取何值，以得 θ 的一個優良推定值。若 $c = 1$，此風險函數有常數值 1，若 $c \neq 1$，風險函數爲有正導數之 θ 的線性函數，因此當 θ 增大時，此函數值亦隨之增大。若除了 θ 必爲正外不限 θ 之值，則 $c = 1$ 在推定量 $d(X) = cX$ 之群中可得一大中取小推定量，此爲該群中對此風險函數得到一有限最大值的唯一推定量。推定量（estimator）一辭習慣上用來表示隨機變數的函數，而推定值（estimate）一辭則表其數值。

1－3 平均風險

若除了樣本 $X = (X_1, \cdots, X_n)$ 之外，尚有 θ 的事前分配可用，則 X 和 θ 皆視爲隨機變數，由於風險函數 $\mathcal{R}(\theta, d)$ 依存於 θ，因此其亦爲一隨機變數。爲衡量決策函數 d 之整體有效性，我們求此風險函數的平均值，而後用此值做爲比較的基準。此一由 θ 之事前分配所求得的期望值，定義了一新函數，稱爲平均風險或貝氏風險，我們以 r 來表示。

定義 3 若 θ 為有密度 $\pi(\theta)$ 的一連續隨機變數，則平均風險為

$$(4) \qquad r(\pi, d) = E\mathcal{R}(\theta, d) = \int \mathcal{R}(\theta, d)\pi(\theta)\, d\theta$$

若 θ 為一不連續隨機變數，則上述積分以對應之加總號代替。

對一已知的決策函數和事前分配，$r(\pi, d)$ 爲一實數，因此如同大中取小之比較般，我們可由此等數值來比較決策函數。若在所有屬於此種函數群 D 之決策函數中，有一決策函數 d 使 $r(\pi, d)$ 最小，則稱其爲在平均風險意味下的最佳決策函數，唯傳統上稱其爲關於事前分配 π 的貝氏決策函數。

定義 4 若一決策函數 d_0 滿足

$$r(\pi, d_0) = \min_{d \in D} r(\pi, d).$$

則稱為關於事前密度π和決策函數群D的貝氏決策函數。

前述定義中已假設，若含有積分，則所處理之函數爲可積，且若含有符號max和min，則所處理之函數具有最大值或最小值。爲使學生能充分認清max, min和sup, inf間的差異，若需要更一般的定義，我們將以sup和inf代替max和min。

爲說明如何求出平均風險，考慮前節中用來說明推定卜松分配之均數的問題。設θ之事前分配爲指數密度$\pi(\theta) = e^{-\theta}, \theta > 0$，則利用(3)中所得的結果，平均風險變成

$$r(\pi, d) = \int_0^\infty [c^2 + \theta(c-1)^2]e^{-\theta}\, d\theta$$
$$= c^2 + (c-1)^2 = 2c^2 - 2c + 1.$$

在形如$d(x) = cx$之決策函數群中求得貝氏決策函數，我們必須取使$r(\pi, d)$爲最小的c值。由基本的微分法可知，若取$c = 1/2$，則$r(\pi, d)$爲最小；因此$d(X) = X/2$爲對應於事前密度$\pi(\theta) = e^{-\theta}$的貝氏推定量。

由此一結果可看出一有趣的事實，即在形如$d(x) = cx$的推定群中，貝氏推定值僅爲大中取小推定值的一半。由於事前密度$\pi(\theta) = e^{-\theta}$對$\theta$之較小值給予的權數遠大於$\theta$之較大值，因此很自然地，其所得的值比沒有用此一情報之推定量$d(X) = X$的值小。

1－4 損失函數的選取

損失函數之選取依問題的性質，和各種可能不正確決策之嚴重程度的評估而定。例如，一實驗者在比較大誤差與中等程度誤差時，可能有相當確定的理念知曉是否其所給予的損失嚴重程度過大。但是，實驗者却很少具有此種理念，因而通常需要統計學家的建議。很幸運的，在很多問題中選取損失函數時可有相當大的變異，而對最適決策函數的性質沒有任何改變，此尤以樣本大小大時爲然。

若大誤差是一非常嚴重的事，我們可取損失函數爲$\mathscr{L}(\theta, a) = c(a-\theta)^2$，此地之$c$爲一正常數。若視損失與誤差大小成比例，則取$\mathscr{L}(\theta, a) = c|a-\theta|$。能適應各種情形的函數爲$\mathscr{L}(\theta, a) = c(\theta)|a-\theta|^b$，此地$b$爲某

正常數，$c(\theta)$ 爲 θ 的某正函數（見習題6）。

由於我們所關心者爲發展一系統性的理論，而非實際應用此一理論的細節，因此我們的敍述將偏限於在統計理論和實務中較常用的最簡單損失函數。此等似是太特定化的損失函數，可略微加以一般化，而不致影響到理論，但是如此一般化却會使符號變得很複雜。

對一檢定假設問題，我們將取

$$\mathscr{L}(\theta, a) = \begin{cases} 0, & \text{若決策正確} \\ 1, & \text{若決策錯誤} \end{cases}$$

對多元決策問題，我們將用同一損失函數來檢定假設。亦即若決策正確，賦予 $\mathscr{L}$ 值 0，不管做出 k -1個錯誤決策中的那一個，若決策錯誤，賦予 $\mathscr{L}$ 值 1。

對推定問題，我們將取

$$\mathscr{L}(\theta, a) = (a - \theta)^2.$$

此即衆所週知之二次損失或平方誤差損失。由於在推定問題中取行動 a 即決定 θ 的值爲數 $d(x_1, \cdots, x_n)$，故此地之損失函數通常寫成

$$\mathscr{L}(\theta, d) = (d - \theta)^2,$$

在此 $d = d(x_1, \cdots, x_n)$ 爲 θ 的推定值。

以下各章將先研究推定理論，而後爲檢定假設理論。由於多元決策理論相當困難，因此對它的敍述相當少，該理論之較簡單部份於假設檢定一章之貝氏一節中討論。

習 題

1. 已知 $f(x \mid \theta) = \dfrac{e^{-(1/2)(x-\theta)^2}}{\sqrt{2\pi}}$ 且 $\mathscr{L}(\theta, a) = (a-\theta)^2$，並設已取一 X 的單一觀察值。若取 d 爲函數 $d(x) = cx$，c 爲常數，求 $\mathscr{R}(\theta, d)$ 之值。
2. 對習題 1，在函數 $d(x) = cx$ 群中是否有大中取小決策函數？
3. 已知 $f(x \mid \theta) = \binom{2}{x} \theta^x (1-\theta)^{2-x}$，$x=0,1,2$，$0 < \theta < 1$，且 $\mathscr{L}(\theta, a) = (a-\theta)^2$，