Chapter 9

區間估計

前章點估計所討論的內容中，主要是探討如何求取參數θ之最佳估計式。接下來本章將探討如何利用最佳估計式之抽樣分配來發展關於參數θ的區間估計（interval estimation）。

9.1 區間估計觀念

一、區間估計之觀念

定義1

設$(X_1,X_2,\ldots,X_n)$為由參數θ之母體$f(x;\theta)$抽出之一組隨機樣本，令$L(X_1,X_2,\ldots,X_n)$及$U(X_1,X_2,\ldots,X_n)$為兩個統計量，且使得

$$P\big(L(X_1,X_2,\ldots,X_n)\le\theta\le U(X_1,X_2,\ldots,X_n)\big)=1-\alpha$$

則區間

$$\big(L(X_1,X_2,\ldots,X_n),U(X_1,X_2,\ldots,X_n)\big)$$

稱為參數θ之$100(1-\alpha)\%$的隨機區間（random interval）或稱區間估計量（interval estimator），其中$1-\alpha$稱為信賴係數或信賴度。但當我們獲取確定之一組樣本資料$(x_1,x_2,\ldots,x_n)$代入上述隨機區間時，則區間

$$\big(L(x_1,x_2,\ldots,x_n),U(x_1,x_2,\ldots,x_n)\big)$$

稱為參數θ之$100(1-\alpha)\%$的信賴區間（confidence interval）或區間估計值（interval estimate）。

定義2

1.若 $P\left(L(X_1,X_2,\ldots,X_n)\le\theta\right)=1-\alpha$ ，則區間

$\left(L(X_1,X_2,\ldots,X_n),\infty\right)$

稱為參數θ之$100(1-\alpha)\%$之下單尾（one-sided lower）隨機區間。

2.若 $P\left(\theta\le U(X_1,X_2,\ldots,X_n)\right)=1-\alpha$ ，則區間

$\left(-\infty,U(X_1,X_2,\ldots,X_n)\right)$

稱為參數θ之$100(1-\alpha)\%$之上單尾（one-sided upper）隨機區間。

定義3

設$\hat{\Theta}$為參數θ之估計式，若$P(|\hat{\Theta}-\theta|\le d)=1-\alpha$則$d$稱為以$\hat{\Theta}$估計$\theta$的$100(1-\alpha)\%$之誤差界限。

二、區間估計推導

通常要尋找某個參數θ之區間估計量，必須先尋找樞紐量（pivotal quantity）然後利用樞紐量直接移項求得區間估計量，但一般樞紐量並不容易尋找，但若針對常見參數μ、p、σ^2等參數之樞紐量就較易尋求，茲先給予樞紐量一個定義。

定義4

若$Q=q(X_1,X_2,\ldots,X_n;\theta)$為一隨機變數，其僅為$X_1,X_2,\ldots,X_n$與參數$\theta$之函數，其中$\theta$為唯一的未知量，且$Q$之機率分配與參數$\theta$或其他任何未知參數無關，則$Q=q(X_1,X_2,\ldots,X_n;\theta)$即稱之為樞紐量。

例如，$T=\dfrac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$為$(X_1,X_2,...,X_n)$與參數$\mu$之函數，其中$\mu$

是唯一的未知量，且 T 之機率分配函數與 μ 無關，故此 $T=\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$ 即爲一個樞紐量。事實上，找到樞紐量後，即可利用前面所提過之各種抽樣分配，諸如 Z 、 T 、 χ^2 、 F 分配之觀念，來推導信賴區間。爲了方便讀者尋找信賴區間，茲將參數區間估計的過程及步驟敘述如下：

㈠尋找未知參數 θ 的最佳點估計式：一般較常用來估計未知參數的點估計式之方法爲最大概似法。

㈡決定抽樣樣本大小：抽樣樣本的大小影響抽樣分配的建立，故須先決定抽樣樣本的大小。

㈢決定點估計式的抽樣分配：可根據第七章抽樣分配內容尋找。

㈣根據點估計式之抽樣分配的觀念來尋找樞紐量，然後再利用此樞紐量來發展信賴區間。

9.2 母體平均數 μ 之估計

母體平均數 μ 之估計是最常見估計問題之一，因樣本平均數 $\bar{X}$ 爲參數 μ 之最佳估計式，故可利用 $\bar{X}$ 之抽樣分配來發展 μ 之估計問題。

一、估計誤差及樣本數問題

由第七章可知，當 $(X_1,X_2,\ldots,X_n)$ 爲抽自下列條件的一組隨機樣本，即

㈠抽自非常態母體，且樣本數 $n>30$ ，或

㈡抽自常態母體，則不論樣本數 n 之大小

則樣本平均數 $\bar{X}$ 服從常態分配，即

$$\bar{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$$

因此由定義3可知

$$P(|\bar{X}-\mu|\le d)=1-\alpha\Rightarrow P\left(|Z|\le\frac{d}{\sigma/\sqrt{n}}\right)=1-\alpha$$