# 8.3 尋找估計式之方法

在8.2節中提過評斷優良估計式之一些準則，但這些優良估計式要如何尋找？便是本節要討論的重點，一般尋找估計式常見方法有：

㈠最大概似法（method of maximum likelihood）。

㈡動差法（method of moment）。

㈢最小平方法（method of least square）。

㈣最佳線性不偏估計法（*BLUE*）。

㈤貝氏估計法（method of Bayes）。

本節除了最小平方法留在迴歸分析時再討論，其餘方法分述如下。

## 一、最大概似法

㈠**最大概似法之觀念**：一般母體之參數 $\theta$ 皆未知，今若從此母體抽出一組隨機樣本，此組樣本之可能性無法得知，因此若能找到一個估計值 $\hat{\theta}$，且可使這組樣本發生之可能性為最大時，則此估計值 $\hat{\theta}$ 即稱為 $\theta$ 之最大概似估計值。

㈡**最大概似估計式之相關定義**：

**定義㈦**：設 $(X_1, X_2, ..., X_n)$ 為抽自母體 $f(x;\theta)$ 之一組隨機樣本，則其概似函數（likelihood function）即為此 $n$ 個隨機變數 $(X_1, X_2, ..., X_n)$ 的聯合機率分配 $f(x_1, x_2, ..., x_n;\theta)$。因為參數 $\theta$ 未知，故此概似函數為 $\theta$ 的函數，一般常將其寫為

$$L(\theta) = f(x_1, x_2, ..., x_n;\theta) = f(x_1;\theta)\cdot f(x_2;\theta)\cdots\cdots f(x_n;\theta)$$
$$= \prod_{i=1}^{n} f(x_i;\theta)$$

**定義㈧**：若 $L(\theta)$ 為概似函數，今有一個估計式 $\hat{\Theta}(X_1, X_2, ..., X_n)$ 可使 $L(\theta)$ 為最大時，則此估計式 $\hat{\Theta}(X_1, X_2, ..., X_n)$ 即稱為參數 $\theta$ 的最大概似估計式（*MLE*）。而當獲取樣本資料值 $(x_1, x_2, ..., x_n)$ 代入上式估計式 $\hat{\Theta}(X_1, X_2, ..., X_n)$ 可知 $\hat{\theta}$，而 $\hat{\theta}$ 即為參數 $\theta$ 之最大概似估計值。

㈢**求最大概似估計式之步驟**：一般在求算 $MLE$ 時，要先瞭解概似函數 $L(\theta)$ 是否可微分，若函數 $L(\theta)$ 可微分則依微積分求極大值之方法，即可快速求出 $MLE$，但若此函數 $L(\theta)$ 不可微分，則必需利用數值分析方法，才可求出 $MLE$。底下將先針對可微分部分來探討。

**定理**㈤：若概似函數 $L(\theta)$ 可微分，則一般在求算 $MLE$ 之步驟如下：

1. 先找概似函數，即 $L(\theta)=f(x_1,x_2,...,x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta)$。
2. 令 $\dfrac{d\ln L(\theta)}{d\theta}=0$，解 $\theta$，可獲得 $\hat{\theta}$。
3. 再檢查 $\left.\dfrac{d^2\ln L(\theta)}{d\theta^2}\right|_{\hat{\theta}}<0$。

則此估計式 $\hat{\Theta}$ 即爲 $\theta$ 之最大概似估計式。

**例題 13**

設隨機變數X為幾何分布 $f_X(x)=p(1-p)^{x-1}$，$x=1,2,\ldots$ 且 $p\in(0,1)$ 自此母體隨機抽取 $X_1$, $X_2,\ldots,X_n$ 之隨機樣本。試以最大概似法（Maximum likelihood estimator）估計 $p$，記為 $\tilde{p}$。　（97國安局三等、98關務）

**Sol**：

因 $L(p)=f(x_1;p)\cdot\cdots\cdot f(x_n;p)=p(1-p)^{x_1-1}\cdot\cdots\cdot p(1-p)^{x_n-1}$

$$=p^n(1-p)^{\sum_{i=1}^{n}x_i-n}$$

又 $\ln L(p)=n\ln p+(\sum_{i=1}^{n}x_i-n)\ln(1-p)$

且 $\dfrac{d\ln L(p)}{dp}=\dfrac{n}{p}+\dfrac{\sum_{i=1}^{n}x_i-n}{1-p}(-1)$；令 $\dfrac{d\ln L(p)}{dp}=0$，則可知

$$\frac{n}{p}=\frac{\sum_{i=1}^{n}x_i-n}{1-p}\Rightarrow n-np=p\sum_{i=1}^{n}x_i-np$$

$$\Rightarrow p=\frac{n}{\sum_{i=1}^{n}x_i}=\frac{1}{\overline{x}}\text{，又}\frac{d^2\ln L(p)}{dp^2}=\frac{-n}{p^2}+\frac{-(\sum_{i=1}^{n}x_i-n)}{(1-p)^2}$$

且 $\left.\frac{d^2\ln L(p)}{dp^2}\right|_{P=\frac{1}{\overline{X}}}=\frac{-n\overline{x}^3}{\overline{x}-1}<0$

故知參數 $p$ 之最概似估計式為 $\hat{P}=\dfrac{1}{\overline{X}}$

**例題 14**

自母數為 $\lambda$ 之卜瓦松分配（Poisson distribution）抽出一大小為 $n$ 的隨機樣本 $X_1,X_2,\ldots,X_n$，以估計此一未知母數 $\lambda$。

⑴試求 $\lambda$ 之最概估計量（Maximum likelihood estimator）$\hat{\lambda}$。

⑵驗證此最概估計量 $\hat{\lambda}$ 是否為 $\lambda$ 之一不偏估計量（Unbiased estimator）。 （94關務三等、99地方政府經行）

**Sol**：

⑴因 $L(\lambda)=f(x_1;\lambda)\cdot\cdots\cdot f(x_n;\lambda)$

$$=\frac{e^{-\lambda}\lambda^{x_1}}{x_1!}\cdot\frac{e^{-\lambda}\lambda^{x_2}}{x_2!}\cdot\cdots\cdot\frac{e^{-\lambda}\lambda^{x_n}}{x_n!}=\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i!}$$

又 $\ln L(\lambda)=-n\lambda+\left(\sum_{i=1}^{n}x_i\right)\ln\lambda-\ln\left(\prod_{i=1}^{n}x_i!\right)$，且

$\frac{d\ln L(\lambda)}{d\lambda}=-n+\frac{\sum_{i=1}^{n}x_i}{\lambda}$，令 $\frac{d\ln L(\lambda)}{d\lambda}=0$，則

$-n+\frac{\sum_{i=1}^{n}x_i}{\lambda}=0$，即 $\lambda=\frac{\sum_{i=1}^{n}x_i}{n}=\overline{x}$

又 $\dfrac{d^2\ln L(\lambda)}{d\lambda^2}=\dfrac{-\sum_{i=1}^{n}x_i}{\lambda^2}$，且 $\left.\dfrac{d^2\ln L(\lambda)}{d\lambda^2}\right|_{\lambda=\bar{x}}=\dfrac{-\sum_{i=1}^{n}x_i}{\bar{x}^2}=\dfrac{-n}{\bar{x}}<0$

故知 $\lambda$ 之最大概似估計式為 $\hat{\lambda}=\bar{X}$。

⑵ $E(\hat{\lambda})=E(\bar{X})=E\left(\dfrac{\sum_{i=1}^{n}X_i}{n}\right)=\dfrac{1}{n}\sum_{i=1}^{n}E(X_i)=\dfrac{1}{n}\sum_{i=1}^{n}\lambda=\dfrac{1}{n}n\lambda=\lambda$

故知 $\hat{\lambda}=\bar{X}$ 為母數入之不偏估計式。

**例題 15**

設 $(X_1,X_2,\ldots,X_n)$ 為抽自 $N(\mu,\sigma^2)$ 之一組隨機樣本，試求 $\mu$ 及 $\sigma^2$ 之最大概似估計量。

**Sol**：

因 $L(\mu,\sigma^2)=\prod_{i=1}^{n}f(x_i;\mu,\sigma^2)=\prod_{i=1}^{n}\dfrac{1}{\sqrt{2\pi}\sigma}exp\left\{-\dfrac{(x_i-\mu)^2}{2\sigma^2}\right\}$

$$=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right\}$$

又 $\ln L(\mu,\sigma^2)=-\dfrac{n}{2}\ln(2\pi)-\dfrac{n}{2}\ln\sigma^2-\dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}$，所以

$$\frac{\partial\ln L(\mu,\sigma^2)}{\partial\mu}=\frac{-2\sum_{i=1}^{n}(x_i-\mu)(-1)}{2\sigma^2}$$

$$\frac{\partial\ln L(\mu,\sigma^2)}{\partial\sigma^2}=-\frac{n}{2\sigma^2}+\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4}$$

又令 $\dfrac{\partial\ln L(\mu,\sigma^2)}{\partial\mu}=0$ 且 $\dfrac{\partial\ln L(\mu,\sigma^2)}{\partial\sigma^2}=0$，即可獲得

$$\begin{cases} \dfrac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2}=0 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (1) \\ \dfrac{-n}{2\sigma^2}+\dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4}=0 \cdots\cdots\cdots\cdots\cdots\cdots (2) \end{cases}$$

再解此聯立方程式，由第(1)式中可知

$\sum_{i=1}^{n}(x_i-\mu)=0$，即 $\sum_{i=1}^{n}x_i-n\mu=0$，所以 $\mu=\overline{x}$

今將 $\mu=\overline{x}$ 代入第(2)式知

$\dfrac{-n}{2\sigma^2}+\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{2\sigma^4}=0$，即 $\sigma^2=\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n}$

故參數 $\mu$ 之最大概似估計式為 $\hat{\mu}=\overline{X}$

而參數 $\sigma^2$ 之最大概似估計式為 $\hat{\sigma}^2=\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n}$

在上述所舉例子中可發現，其求解過程皆有一定程序，但有些概似函數 $L(\theta)$ 並無法以定理㈤之三個步驟求得 $MLE$。此時可能要利用數值分析方式來探討才能獲得參數 $\theta$ 之 $MLE$，且各種不同題型求算方式可能也不同，底下所舉的例子皆爲不可微，因此讀者必須多練習才能獲得解題技巧。

**例題 16**

設隨機變數 $X$ 為分布於 $[0,\theta]$，$\theta>0$ 的均勻分配，自此母體隨機抽取 $X_1$, $X_2,\ldots,X_n$ 之隨機樣本。試以最大概似法（Maximum likelihood estimator）估計 $\theta$，記為 $\hat{\theta}$。（95原住民三等、高考）

**Sol**：

因 $L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\dfrac{1}{\theta}\cdot\dfrac{1}{\theta}\cdot\cdots\cdot\dfrac{1}{\theta}=\dfrac{1}{\theta^n}$

又 $L(\theta)$ 為 $\theta$ 之嚴格遞減函數，今欲使 $L(\theta)$ 為最大時，$\theta$ 值應該取愈小

愈好；又因$0 \le x_i \le \theta$，$i = 1,2,...,n$，且

$$0 \le x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)} \le \theta$$

因此當$\theta$取$\hat{\theta} = x_{(n)}$時能使$L(\theta)$為最大。故知$\theta$之最大概似估計式為

$$\hat{\theta} = X_{(n)} = \max(X_1, X_2, ..., X_n)$$

## 例題 17

設$(X_1, X_2, ..., X_n)$為抽自下列母體之一組隨機樣本

$$f(x) = \frac{1}{2c} \quad , \theta - c \le x \le \theta + c$$

試求參數$\theta$之最大概估計式，記為$\hat{\theta}$。（高考）

**Sol**：

因$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \left[\frac{1}{2c}\right]^n$，$\theta - c \le x_i \le \theta + c$

故$\theta$取任何點皆能使$L(\theta)$最大，但因

$\theta - c \le x_i \le \theta + c$，$i = 1,2,...,n$，且知

$\theta - c \le x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)} \le \theta + c$，又

$$\begin{cases} \theta - c \le x_{(1)} \Rightarrow \theta \le x_{(1)} + c \\ \theta + c \ge x_{(n)} \Rightarrow \theta \ge x_{(n)} - c \end{cases}$$

即$x_{(n)} - c \le \theta \le x_{(1)} + c$

所以$\theta$之最大概似估計式為取介於區間$[X_{(n)} - c, X_{(1)} + c]$中之任一估計式，可以$\hat{\theta} = \alpha(X_{(n)} - c) + (1-\alpha)(X_{(1)} + c)$表示，其中$0 \le \alpha \le 1$。

## 例題 18

在阿姆斯特丹街頭看車流，請利用最大概似法估計該市自行車總數。假設該市發自行車牌照，按流水號數，且自1號發起。

**Sol**：

因任一輛車被看到之機會皆相同，故

$$f(x)=\frac{1}{N}\quad ，x=1,2,...,N$$

因 $L(N)=\prod_{i=1}^{n}f(x_i;N)=\frac{1}{N}\cdot\frac{1}{N}\cdot\cdots\cdot\frac{1}{N}=\frac{1}{N^n}$

又 $L(N)$ 為 $N$ 之嚴格遞減函數，故當 $N$ 愈小時，$L(N)$ 愈大，

又因 $1\le x_i\le N$ ，$i=1,2,...,n$

且 $1\le x_{(1)}\le x_{(2)}\le\cdots\le x_{(n)}\le N$ ，所以當 $N$ 取最小 $x_{(n)}$ 時 $L(N)$ 為最大，亦即 $N$ 之 $MLE$ 為 $\hat{N}=X_{(n)}=\max(X_1,...,X_n)$

㈣ *MLE*之性質：

1. $MLE$ 不一定具有不偏性。
2. $MLE$ 爲在最樂觀的評判標準下所選出之估計式，故並不能保證爲最好的估計式。
3. $MLE$ 不一定只有一個，即不唯一。
4. $MLE(\hat{\Theta})\overset{\text{漸近}}{\sim}N(\theta,CRLB)$ 。
5. $MLE$ 具有不變性（invariance），即若 $\hat{\Theta}$ 爲 $\theta$ 的最大概似估計式，則 $u(\hat{\Theta})$ 亦爲 $u(\theta)$ 之最大概似估計式，其中 $u(\theta)$ 爲 $\theta$ 之任意函數。

**例題 19**

在例題14中，試求：

$P(X=0)$ 及 $P(X\ge 1)$ 之最大概似估計式。

**Sol**：

由例題14知 $\lambda$ 之$MLE$為 $\hat{\lambda}=\overline{X}$ ，又

$$P(X=0)=\frac{e^{-\lambda}\cdot\lambda^0}{0!}=e^{-\lambda}$$

故由$MLE$之不變性知，$P(X=0)=e^{-\lambda}$ 之$MLE$為 $e^{-\overline{X}}$

又 $P(X\ge 1)=1-P(X=0)=1-\frac{e^{-\lambda}\cdot\lambda^0}{0!}=1-e^{-\lambda}$

故由$MLE$之不變性知，$P(X\ge 1)=1-e^{-\lambda}$ 之$MLE$為 $1-e^{-\overline{X}}$

## 二、動差法

㈠**動差法之觀念**：動差法是因爲隨機樣本 $(X_1, X_2, ..., X_n)$ 的 $k$ 階原動差，爲母體 $f(x)$ 的 $k$ 階原動差之不偏估計式，因此利用此二種原動差近似相等的觀念，所建立之一種求估計式之方法稱之爲動差法。

㈡**動差法之定義**：

**定義㈨**：設 $(X_1, X_2, ..., X_n)$ 爲抽自母體 $f(x;\theta)$ 之一組隨機樣本，則利用樣本 $k$ 階原動差，去估計母體 $k$ 階原動差而獲得母體參數 $\theta$ 之估計式，即爲動差估計式（ $MME$ ），而此方法稱爲動差法。

㈢**動差法之求解過程**：令 $\mu_k' = m_k'$ ， $k = 1, 2, ..., p$ ，此處 $p$ 表示母體中參數的個數，再求解這組聯立方程式，便可解出參數之動差估計式。其中

$\mu_k' = E(X^k)$ 爲隨機變數對原點的 $k$ 階動差。

$m_k' = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k$ 爲 $(X_1, X_2, ..., X_n)$ 對原點的 $k$ 階樣本動差。

**例題 20**

試求例16中 $\theta$ 之動差估計式 $\hat{\theta}$ ，並驗證 $\hat{\theta}$ 是否具有不偏性。

（95原住民三等）

**Sol**：

因 $X \sim U(0, \theta)$ ，故知 $X$ 之一階原動差為

$$\mu_1' = E(X) = \frac{\theta}{2}$$

又樣本一階原動差為

$$m_1' = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

今令 $\mu_1' = m_1' \Rightarrow \dfrac{\theta}{2} = \bar{x} \Rightarrow \theta = 2\bar{x}$

故 $\theta$ 之動差估計式為 $\hat{\theta} = 2\bar{X}$

又 $E(\hat{\theta}) = E(2\bar{X}) = 2E\left(\dfrac{1}{n}\sum_{i=1}^{n} X_i\right) = \dfrac{2}{n}\sum_{i=1}^{n} E(X_i) = \dfrac{2}{n}\sum_{i=1}^{n}\left(\dfrac{\theta}{2}\right) = \dfrac{2}{n}\dfrac{n\theta}{2} = \theta$