第 10 章　統計估計-點估計

林惠玲　陳正倉著 雙葉書廊發行 2004

學習目的

1.了解點估計的意義、估計的步驟與限制。

2.了解優良估計式的性質。

3.了解判斷估計式的注意事項。

4.了解統計估計的方法。

5.利用 Excel做統計估計。

本章結構

- 統計估計—點估計
  - 點估計
    - 點估計的意義
    - 點估計的步驟
    - 點估計的限制
  - 估計式的評斷標準
    - 不偏性
    - 有效性
      - 絕對有效性
      - 相對有效性
    - 最小變異不偏性
    - 漸近不偏性
    - 一致性
    - 充分性
  - 判斷估計式的應注意事項
  - 統計(點)估計的方法
    - 最小平方法
    - 動差法
    - 最大概似法
  - Excel的使用

- Concept

推論統計學: 利用樣本統計量及其抽樣分配來對母體參數進行推估, 以瞭解母體的特性.

母數統計學 (parametric statistics): 假設母體爲常態或樣本爲大樣本的情況下的推論統計學

無母數統計學(nonparametric statistics): 指母體分配不是常態或樣本爲小樣本的情況下的推論統計學

貝氏統計學 (Baysian statistics): 利用樣本訊息去修正母體參數的事後機率分配, 再利用該事後機率分配去推論母體參數.

※目的

1. 統計估計 (statistical estimation)
2. 假設檢定 (hypothesis testing)

統計估計

統計估計

統計估計是利用樣本統計量去推估母體參數的方法。

EX10.1　2007年12月21日的民意調查結果爲：謝蘇配的支持率爲35%，馬蕭配38%. 以電話後四碼隨機抽樣進行電話訪問，共訪問1443位台灣地區20歲以上公民，在95%的信心水準下，抽樣誤差約爲正負2.6個百分點

→ 意義爲何?

10.1 點估計的意義與限制

10.1.1 點估計的意義

由母體抽取一組樣本數爲n的隨機樣本，並以由此得到的樣本統計量做爲母體參數的估計值。

EX: 由 $\overline{X}$ 來估計μ, 則 $\overline{X}$ 爲一估計式

將樣本觀察值代入, 可得到某一數值. 如 $\overline{X}_0 = 125$,則此一數值125即爲估計值

10.1.2 點估計的步驟

①抽取具代表性的樣本

②選擇一個較佳的樣本統計量做爲估計式

③計算樣本統計量的值

④以樣本統計量的值推論母體參數值並做決策

林惠玲 陳正倉著 雙葉書廊發行 2004

EX10.2 想買一間台北市30~40平的房子

Solution:
1. 隨機抽取36間面積30~40坪的房屋為一組樣本, 價格如下

表10.1　台北市區36間房屋的價格

595	1,400	620	790	800	820	1,250	1,380	1,388
890	898	928	1,100	950	960	980	980	620
650	698	698	750	750	760	998	1,050	1,080
838	850	850	850	850	850	860	880	930

（資料來源：各房屋仲介公司 93 年 2 月）

2. 估計如下

$$\overline{X} = \frac{\sum X_i}{n} = 903.92 \quad \text{萬元}$$

$$S = \sqrt{\frac{\sum (X - \overline{X})^2}{n-1}} = 205 \quad \text{萬元}$$

結論: 台北市面積30~40坪的房屋的平均價格爲903.92萬元, 標準差爲205萬元.

→ 可信嗎?

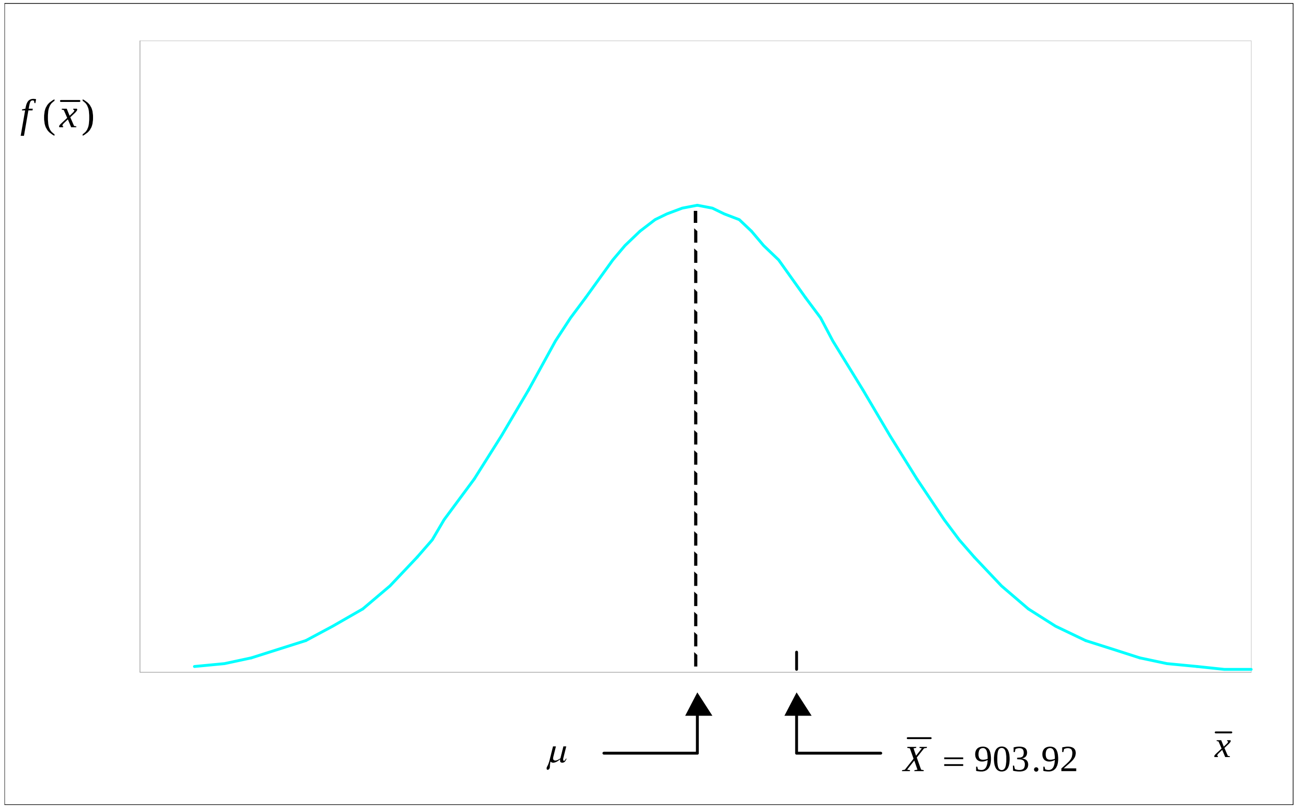
表10.2 台北市區房屋價格的估計

	A	B
1	台北市區房屋價格的估計	
2		
3	平均數	903.92
4	標準誤	34.17
5	標準差	205.01
6	變異數	42030.31
7	個數	36

林惠玲 陳正倉著 雙葉書廊發行 2004

10.1.3 點估計的限制

圖10.1　台北市區房屋價格的估計



10.2 估計式的評斷標準

Q: 什麼是良好的估計式?

1. 不偏性 **(unbiasedness)**
2. 有效性 **(efficiency)**
3. 最小變異不偏性 (minimum variance unbiasedness)
4. 漸近不偏性 (asymptotic unbiasedness)
5. 一致性 **(consistency)**
6. 充分性 (sufficiency)

若符合, 估計值平均而言最接近母體參數的真值

θ: 隨機變數的母體參數， $\hat{\theta}$ 代表θ的估計式

$\hat{\theta} - \theta$ 估計偏誤 (誤差)

$(\hat{\theta} - \theta)^2$ 平方誤差 (損失函數)

估計式的評斷標準

不偏性

若估計式的抽樣分配的平均數等於母體參數值，則該估計式為不偏估計式，否則為偏誤估計式。即若 $E(\hat{\theta}-\theta)=0$ ，則 $\hat{\theta}$ 為 θ 的不偏估計式。

圖10.2 不偏估計式

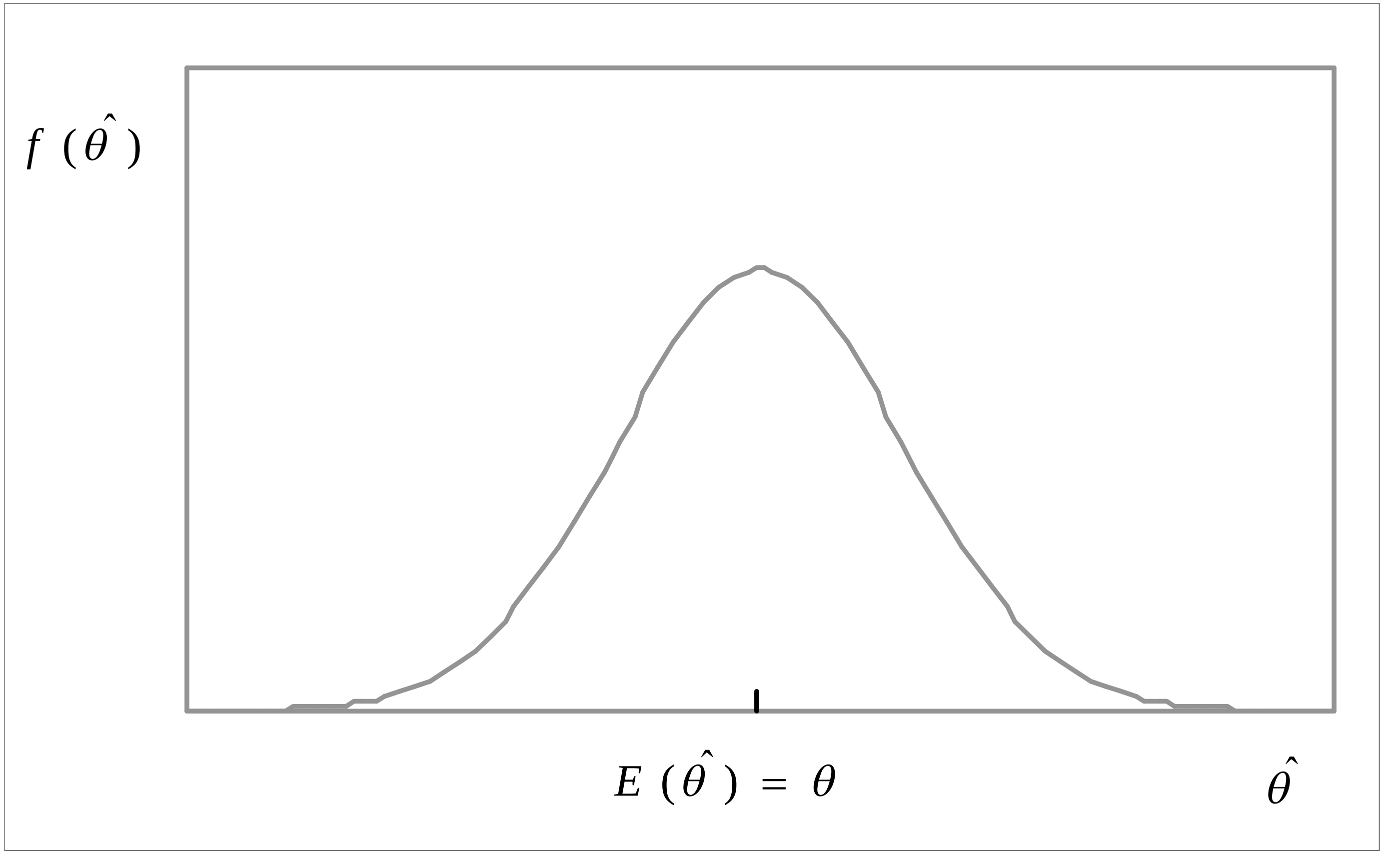


圖10.3 正偏誤估計

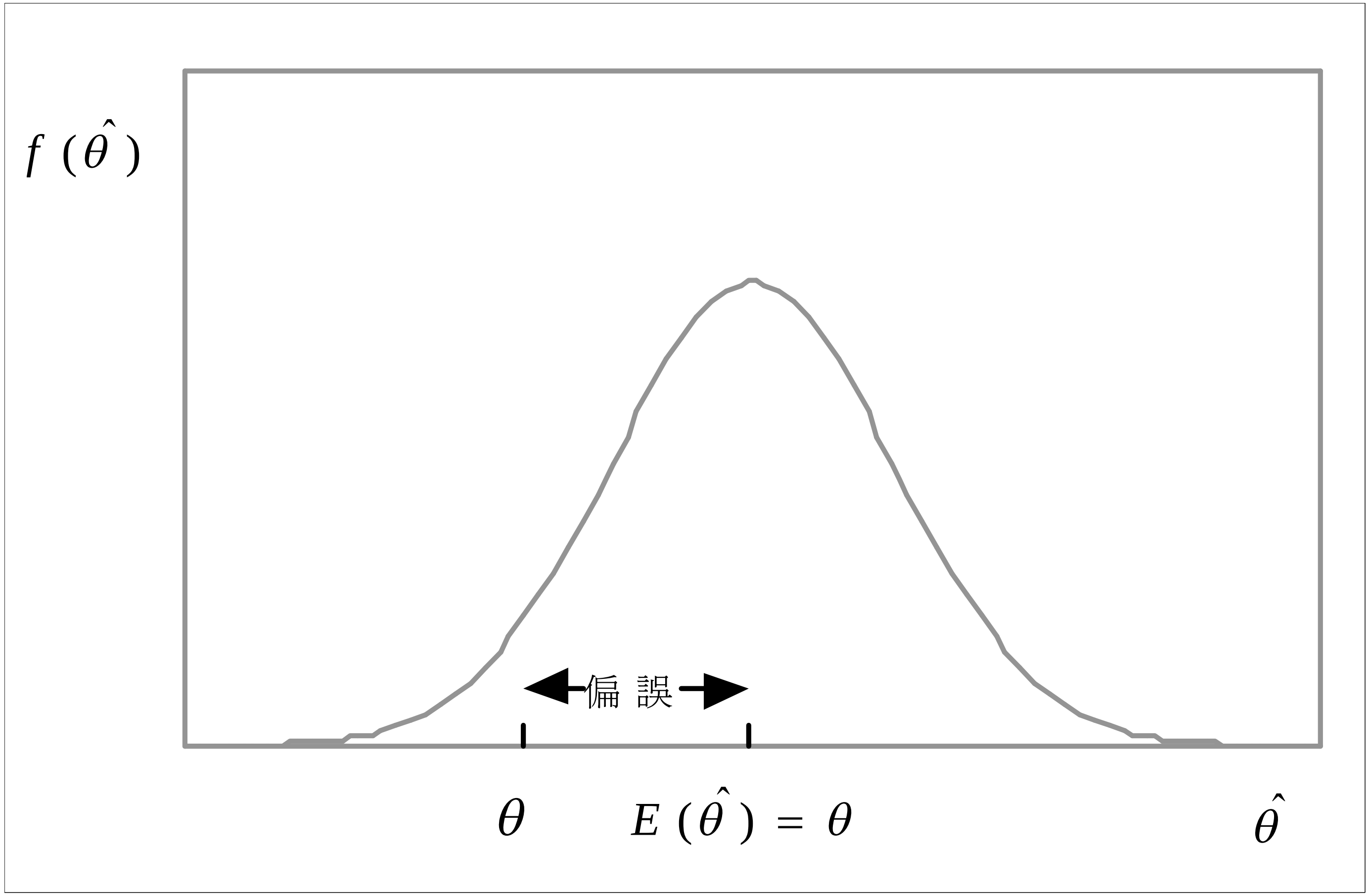
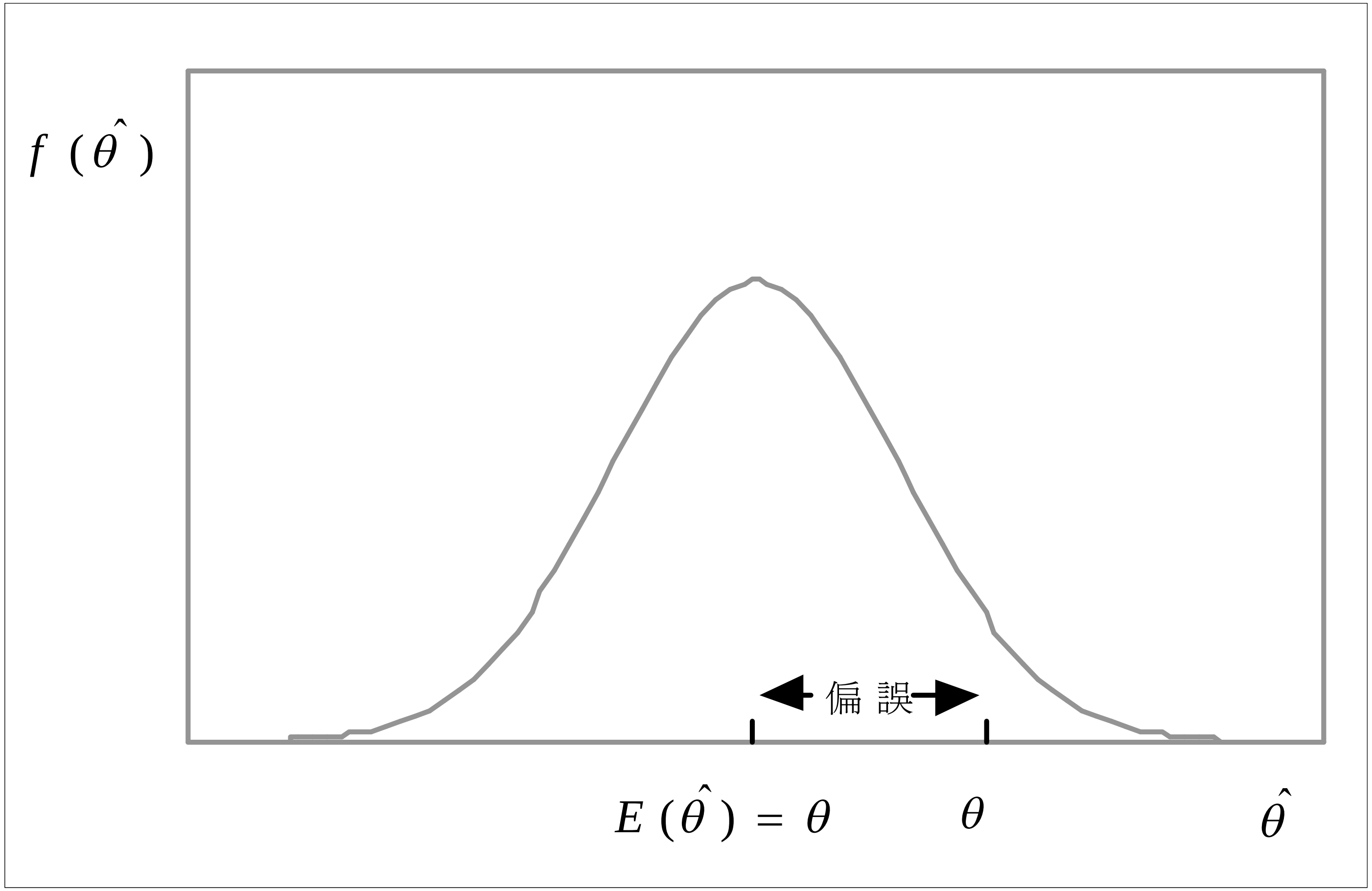


圖10.4 負偏誤估計



EX10.4 楊老師想在某國小邊開一家炸雞店. 該校已做過一次全校性調查(母體)如表10.3. 楊老師當然無該資料, 自己再隨機抽取100人進行抽樣調查, 結果如表10.4

1. 小朋友一星期吃漢堡炸雞的平均數μ(母體平均數)及樣本平均數 $\overline{X}$ 各爲何?
2. 若以樣本平均數估計母體平均數, 估計誤差為何? 偏誤為多少?

表10.3 吃漢堡的機率分配（母體）

x	$f(x)$
0	0.4
1	0.2
2	0.3
3	0.1

林惠玲 陳正倉著 雙葉書廊發行 2004

表10.4　吃漢堡的相對次數分配（樣本）

x	相對次數
0	0.42
1	0.18
2	0.35
3	0.05

Solution:

1. 母體平均數

μ= 0*(0.4)+1*(0.2)+2*(0.3)+3*(0.1)=1.1

樣本平均數

$$\overline{X} = 0\times(0.42)+1\times(0.18)+2\times(0.30)+3\times(0.1)=1.03$$

2. 估計誤差爲

$$\overline{X}-\mu = 1.03-1.1=-0.07$$

但偏誤爲 $E(\overline{X})-\mu = 0$

就單一樣本而言, 估計誤差爲-0.07, 但就所有樣本而言, 平均估計誤差爲0, 因 $\overline{X}$ 爲不偏估計量

林惠玲 陳正倉著 雙葉書廊發行 2004

EX10.6 令 $\hat{\mu}=\dfrac{\sum X}{n+1}$ 為μ的一個估計式, 問 $\hat{\mu}$ 為μ的不偏估計式嗎?

Solution:

$$E(\hat{\mu})=E(\frac{\sum X}{n+1})=\frac{1}{n+1}E\left(\sum X\right)=\frac{n}{n+1}\mu$$

而 $\hat{\mu}$ 的平均估計偏誤可表為

$$Bias(\hat{\mu})=E(\hat{\mu})-\mu=\frac{n}{n+1}\mu-\mu=-\frac{1}{n+1}\mu$$

故 $\hat{\mu}$ 為μ的低估估計式, 不是母體平均數的不偏估計式

EX 10.7 樣本變異數S^2是母體變異數σ^2的不偏估計式嗎?

Solution:

樣本變異數爲 $$S^2=\frac{\sum(X-\bar{X})^2}{n-1}$$

而 $$E(S^2)=E\left[\frac{\sum(X-\bar{X})^2}{n-1}\right]$$

$$=E\left[\frac{\sum X^2-n\bar{X}^2}{n-1}\right]=\frac{1}{n-1}\left[\sum E(X^2)-nE(\bar{X}^2)\right]$$

$$=\frac{1}{n-1}\left[n(\mu^2+\sigma^2)-n\left(\frac{\sigma^2}{n}+\mu^2\right)\right]$$

$$=\frac{1}{n-1}\left[(n-1)\sigma^2\right]=\sigma^2$$

因 1. $E(X^2)=\sigma^2+\mu^2$ 2. $E(\bar{X}^2)=\sigma_{\bar{X}}^2+\mu_{\bar{X}}^2=\dfrac{\sigma^2}{n}+\mu^2$

表10.5　幾個常用的估計式的不偏性

估計式	不偏性
$\overline{X}$ 估計 μ	不偏
S^2 估計 σ^2（樣本抽出放回或母體無限）	不偏
S 估計 σ	偏誤
$\hat{p}$ 估計 p	不偏
$\hat{p}\hat{q}$ 估計 pq	偏誤
$\overline{X}_1 - \overline{X}_2$ 估計 $\mu_1 - \mu_2$	不偏
$\hat{p}_1 - \hat{p}_2$ 估計 $p_1 - p_2$	不偏

NOTE:

1. 不偏性是指估計式所有可能的值的*平均數*等於母體參數, 但就單一數值而言, 仍可能有估計偏誤.
2. 一般而言,不偏估計式較偏誤估計式爲佳, 且較方便用來進行區間估計與假設檢定的推論估計. (尚未涉及損失函數的問題)
3. 若一估計式的偏誤不大,且合乎其他的優良特性時,則仍不失爲一個的估計式.
4. 不偏估計式可能找不到,也可能很多個.

EX μ的不偏估計式 有多個, 可表爲

$$\frac{\sum a_i X_i}{\sum a_i}$$

- 不偏性的定理

若 $\hat{\theta}_1$ 為 θ_1 的不偏估計式, $\hat{\theta}_2$ 為 θ_2 的不偏估計式, 則 $a\hat{\theta}_1 + b\hat{\theta}_2$ (a, b為常數)必為 $a\theta_1 + b\theta_2$ 的不偏估計式

EX10.9 從平均數μ,變異數 σ^2 的母體中, 隨機抽取n個樣本 (n>4) X1, X2, …,Xn, 若

$$T_1 = \frac{X_1 + 2X_2 + X_3}{4}$$

$$T_2 = \frac{1}{4}X_1 + \frac{X_2 + \ldots + X_{n-1}}{2(n-2)} + \frac{1}{4}X_n$$

$$T_3 = \frac{X_1 + \ldots + X_n}{n} = \overline{X}$$

為μ的三個估計式, 問此三個估計式是否皆為不偏?

林惠玲 陳正倉著 雙葉書廊發行 2004

Solution:

$$E(T_1) = E(\frac{X_1 + 2X_2 + X_3}{4})$$
$$= \frac{1}{4}E(X_1) + \frac{2}{4}E(X_2) + \frac{1}{4}E(X_3)$$
$$= \frac{1}{4}\mu + \frac{2}{4}\mu + \frac{1}{4}\mu = \mu$$

$$E(T_3) = E(\frac{\sum X}{n}) = \frac{1}{n}E(\sum X) = \mu$$

$$E(T_2) = E(\frac{1}{4}X_1 + \frac{X_2 + \ldots + X_{n-1}}{2(n-2)} + \frac{1}{4}X_n)$$
$$= \frac{1}{4}E(X_1) + \frac{E(X_2) + \ldots + E(X_{n-1})}{2(n-2)} + \frac{1}{4}E(X_n)$$
$$= \frac{1}{4}\mu + \frac{(n-2)\mu}{2(n-2)} + \frac{1}{4}\mu = \mu$$

林惠玲 陳正倉著 雙葉書廊發行 2004

10.2.2 有效性

有效性是以估計式的**平均平方誤差**來衡量，越小代表估計式的有效性越高

分為: 絕對有效性(absolute efficiency)

相對有效性 (relative efficiency)

❍ 絕對有效性

設$\hat{\theta}$爲θ之估計式，若$\hat{\theta}$的平均平方誤差$MSE(\hat{\theta})=E(\hat{\theta}-\theta)^2$，爲所有估計式中最小者，則稱$\hat{\theta}$在估計$\theta$時具絕對有效性。

因具絕對有效性的估計式不易

相對有效性

設 $\hat{\theta}$ 、$\hat{\hat{\theta}}$ 均爲 θ 的不偏誤估計式，若 $\hat{\theta}$ 的平均平方誤差相對 $\hat{\hat{\theta}}$ 的平均平方誤差較小,即

$$MSE\ (\hat{\theta})\ /\ MSE\ (\hat{\hat{\theta}}) < 1$$

則稱 $\hat{\theta}$ 相對 $\hat{\hat{\theta}}$ 爲在估計 θ 時具相對有效性。

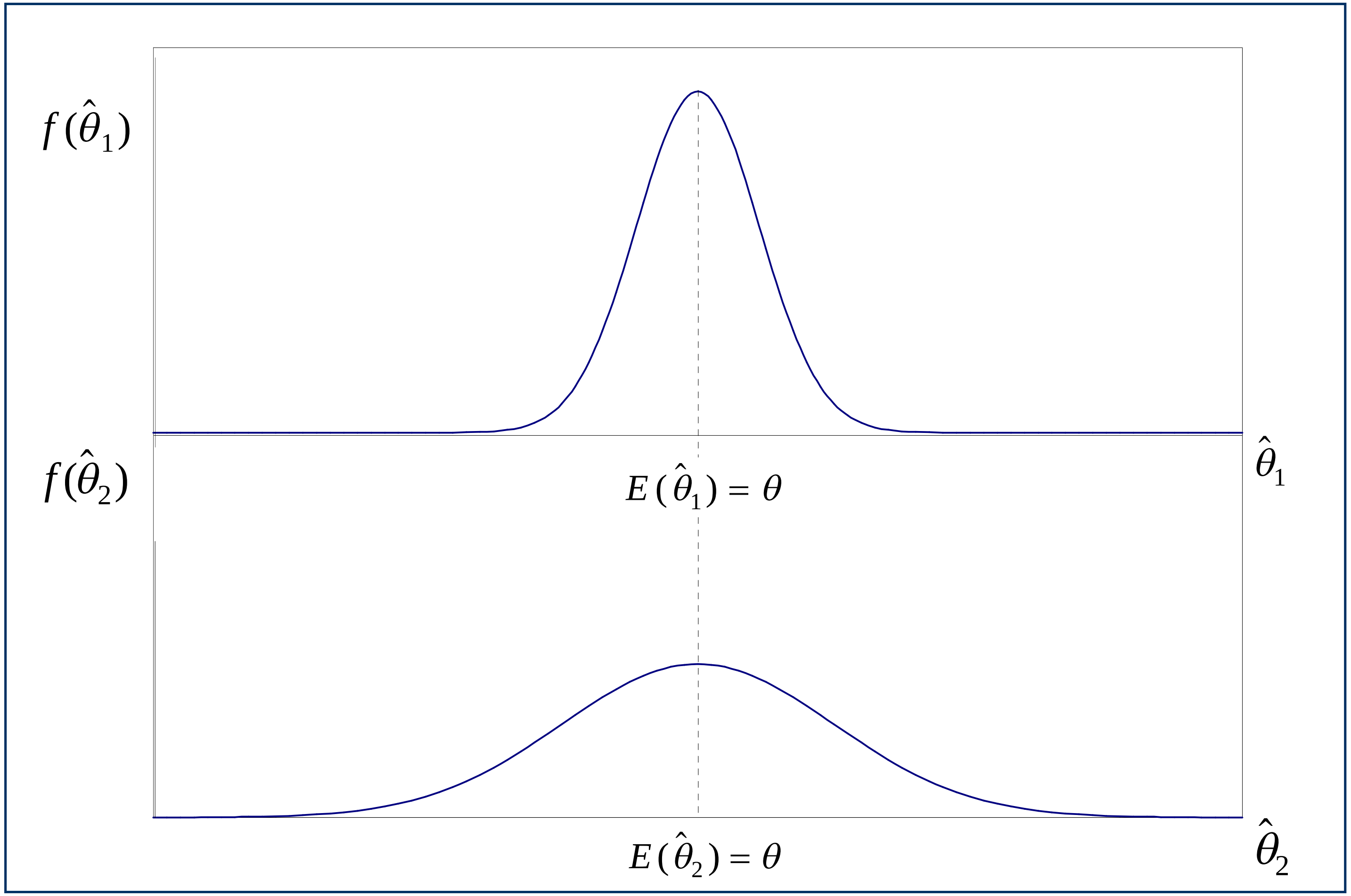
不偏估計式的相對有效性

設 $\hat{\theta}$ 、$\hat{\hat{\theta}}$ 均爲 θ 的不偏誤估計式，若 $V(\hat{\theta})\ /\ V(\hat{\hat{\theta}}) < 1$

則 $\hat{\theta}$ 相對 $\hat{\hat{\theta}}$ 爲有效估計式

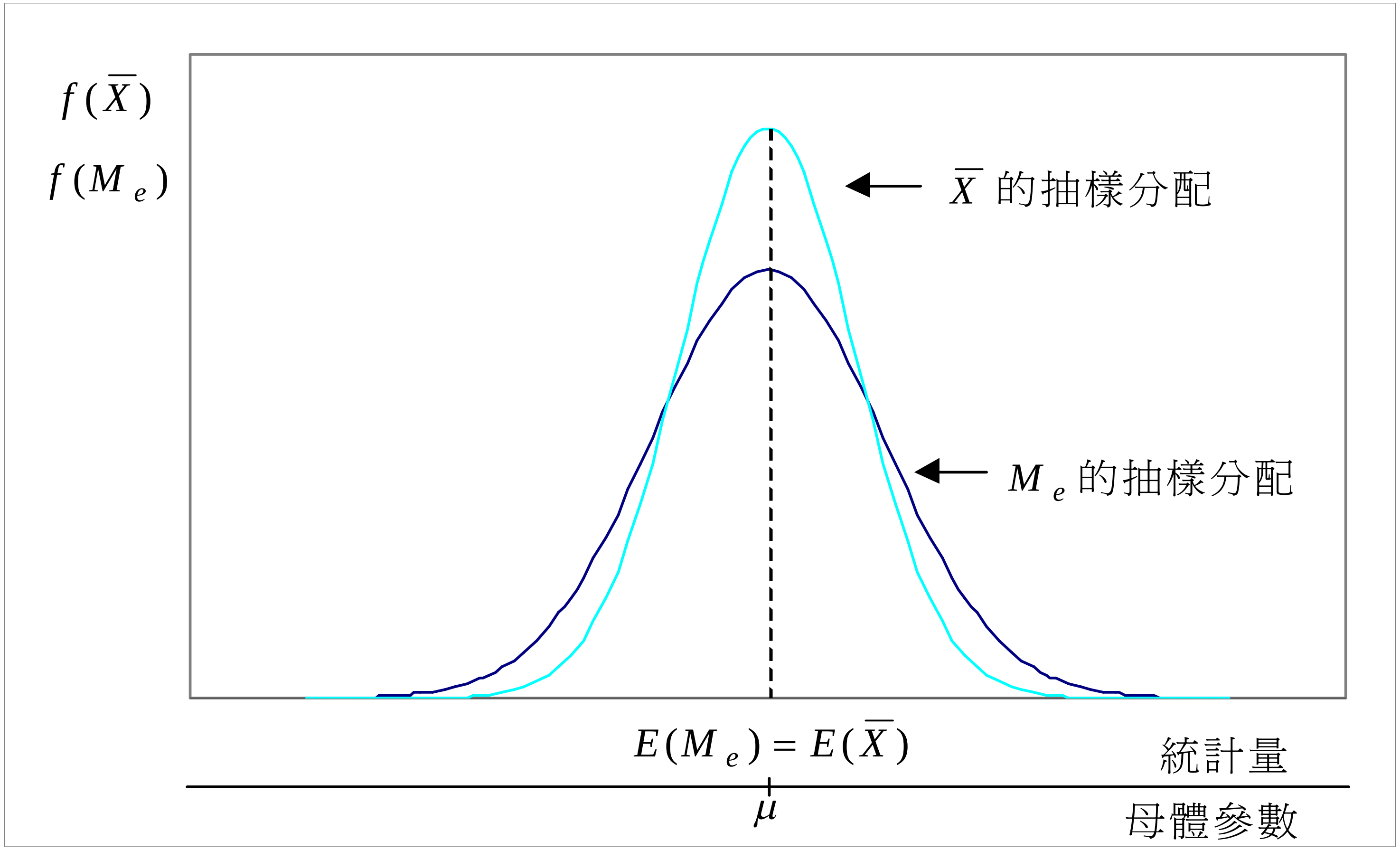
林惠玲 陳正倉著 雙葉書廊發行 2004

圖10.5 估計式的相對有效性



EX10.10

圖10.6 $\overline{X}$ 與 m_e 的相對有效性



EX10.11 令 $\hat{\mu}=\dfrac{\sum X}{n-1}$ 爲母體參數μ的估計式, 試比較 $\hat{\mu}$ 與 $\overline{X}$ 在估計μ時之相對有效性

Solution:

$$MSE(\overline{X})=V(\overline{X})+[Bias(\overline{X})]^2=V(\overline{X})=\frac{\sigma^2}{n}$$

又

$$MSE(\hat{\mu})=V(\hat{\mu})+[Bias(\hat{\mu})]^2=\frac{n\sigma^2}{(n-1)^2}+\frac{\mu^2}{(n-1)^2}$$

$$=\frac{n\sigma^2+\mu^2}{(n-1)^2}\quad (\text{because } E(\hat{\mu})-\mu=\frac{\mu}{n-1}$$

因爲 $$\frac{n\sigma^2+\mu^2}{(n-1)^2}>\frac{n\sigma^2+\mu^2}{n^2}=\frac{\sigma^2}{n}+\frac{\mu^2}{n^2}>\frac{\sigma^2}{n}$$

因此 $\dfrac{MSE(\overline{X})}{MSE(\hat{\mu})}<1$ $\overline{X}$ 相對 $\hat{\mu}$ 在估計μ時具相對有效性

圖10.7 高偏差低變異

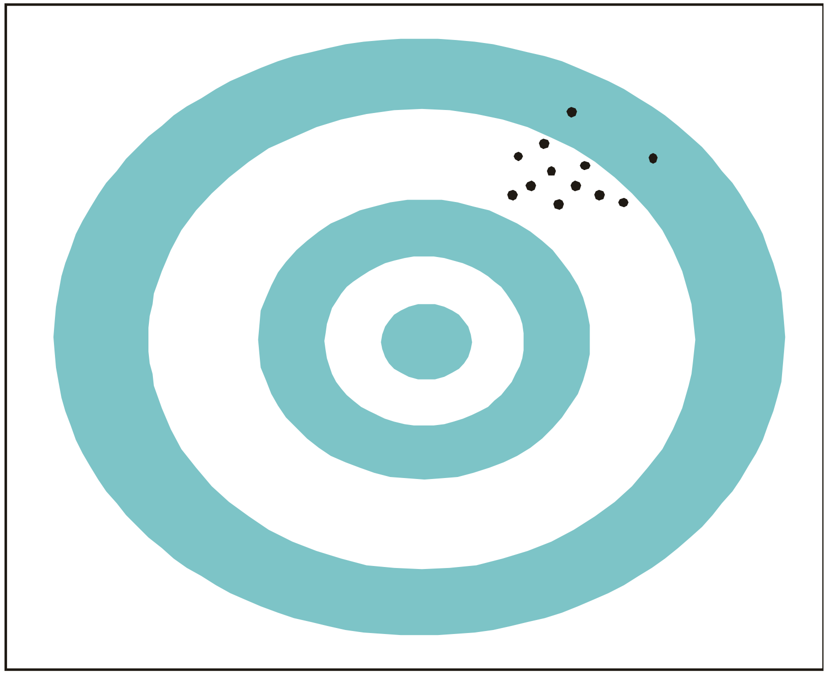


圖10.8 高偏差高變異

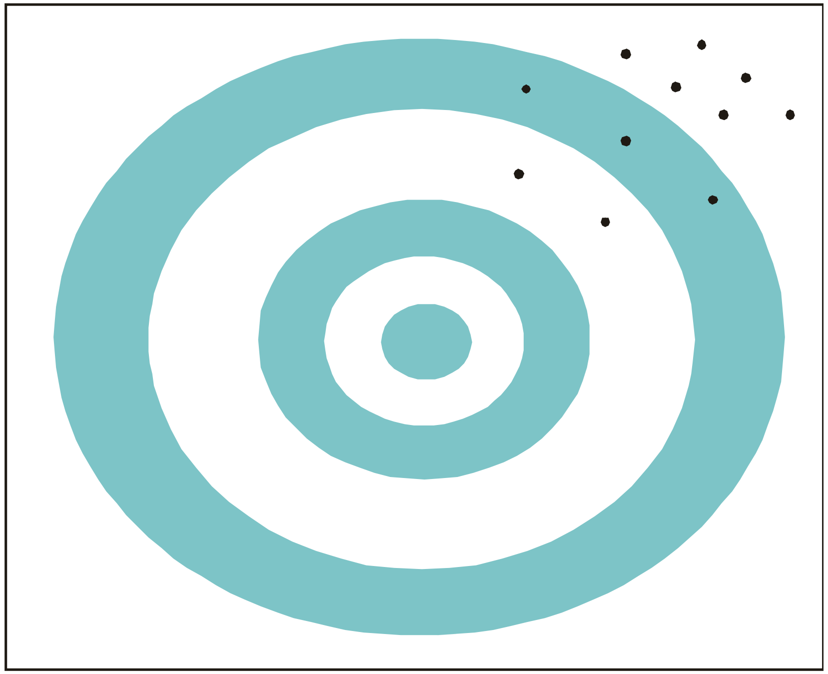


圖10.9　低偏差高變異

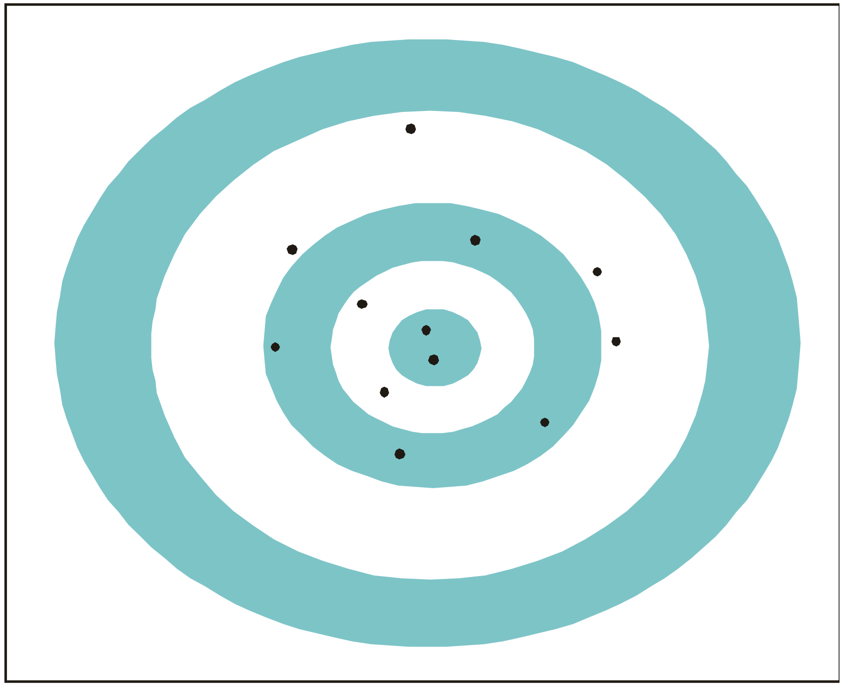
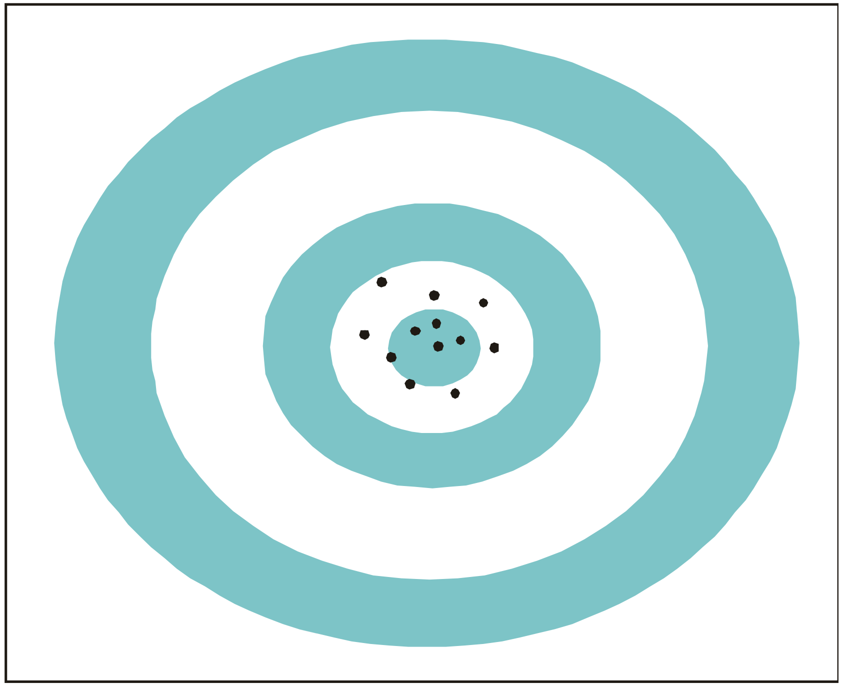


圖10.10　低偏差低變異



林惠玲　陳正倉著 雙葉書廊發行 2004

EX10.13 在例10.9中, 問三個估計式以何者最具有效性

Solution:

$$V(T_1)=V\left(\frac{X_1+2X_2+X_3}{4}\right)=\frac{1}{16}[V(X_1)+4V(X_2)+V(X_3)]=\frac{1}{16}[6\sigma^2]=\frac{3}{8}\sigma^2$$

$$V(T_2)=V\left[\frac{1}{4}X_1+\frac{X_2+\ldots+X_{n-1}}{2(n-2)}+\frac{1}{4}X_n\right]$$

$$=\frac{\sigma^2}{16}+\frac{(n-2)\sigma^2}{4(n-2)^2}+\frac{\sigma^2}{16}=\frac{1}{8}\sigma^2+\frac{\sigma^2}{4(n-2)}$$

$$V(T_3)=V\left[\frac{X_1+\ldots+X_n}{n}\right]=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$$

因爲

$$\frac{V(T_3)}{V(T_1)}=\frac{\sigma^2/n}{\frac{3}{8}\sigma^2}=\frac{8}{3n}<1 \quad \text{if n} > \frac{8}{3}$$

因此　V(T3) < V(T1)

又

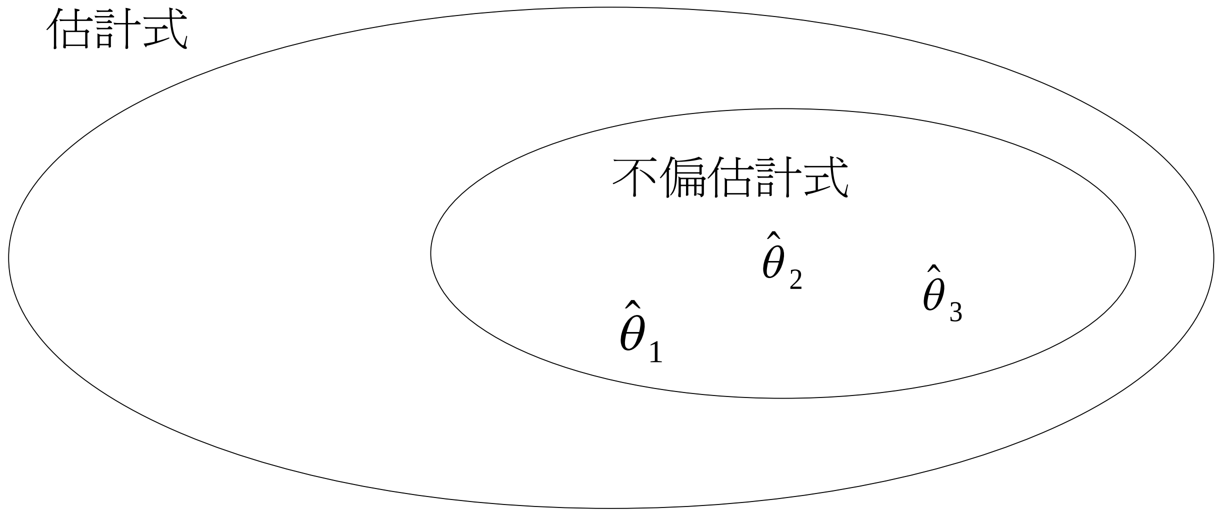
$$V(T_2)-V(T_3)=\frac{1}{8}\sigma^2+\frac{\sigma^2}{4(n-2)}-\frac{\sigma^2}{n}=\left[\frac{(n-4)^2}{8n(n-2)}\right]\sigma^2>0 \quad \text{if n} > 2$$

所以 V(T3) < V(T2)

因此, T3最具有效性

10.2.3 最小變異不偏性

圖10.11 最小變異不偏估計式



$\hat{\theta}_1$: 不偏估計式中變異數最小者。

$V(\hat{\theta}_1) < V(\hat{\theta}_i) \qquad i = 1,2,3\cdots$

EX10.14 設(X1, X2,..,Xn)為來自平均數為μ,變異數為σ^2 的隨機樣本.
令 為Xi的線性估計式(linear estimator),即 $\hat{\mu}=\sum a_i X_i$,其中ai
為常數. 若已知: $\sum a_i = 1$

1. 請證明 $\hat{\mu}$ 為μ的不偏估計式
2. 試求ai使 $\hat{\mu}$ 的變異數最小

Solution:

1.
$$E(\sum a_i X_i)=\sum a_i E(X_i)=\mu\sum a_i=\mu$$

2.
$$V(\sum a_i X_i)=\sum a_i^2 V(X)=\sigma^2\sum a_i^2$$

使 $\sum a_i^2$ 最小, 且滿足 $\sum a_i = 1$, 可解得 $a_i=\frac{1}{n}$

因此 $\hat{\mu}=\frac{1}{n}\sum X_i=\bar{X}$

10.2.4 漸近不偏性

概念

有些估計式雖然在小樣本情況下為偏誤估計式，但當樣本數增大時，偏誤會越來越小而趨近於0，此一性質為極限性質(limiting property)，又稱**大樣本性質**.

漸近不偏性

設$\hat{\theta}_n$爲樣本數爲n時之θ估計式，若$\lim_{n\to\infty} E(\hat{\theta}_n)=\theta$(或$\lim_{n\to\infty} E(\hat{\theta}_n)-\theta=0$)，則$\hat{\theta}_n$稱爲$\theta$的漸近不偏估計式。

林惠玲 陳正倉著 雙葉書廊發行 2004

EX10.15 設 $\hat{\mu}=\dfrac{\sum X}{n+1}$ 為母體參數μ的估計式, 問 $\hat{\mu}$ 為μ的漸近不偏估計式嗎?

Solution:

$$E(\hat{\mu})=E\left(\frac{\sum X}{n+1}\right)=\frac{1}{n+1}E\left(\sum X\right)=\frac{n}{n+1}\mu$$

$$E(\hat{\mu})-\mu=Bias(\hat{\mu})=\frac{-1}{n+1}\mu$$

而 $$\lim_{n\to\infty}\frac{-1}{n+1}\mu=0$$

為漸近不偏

10.2.5 一致性

大樣本性質, 指當樣本數增大時, 估計值會趨近於母體參數真值的機率高

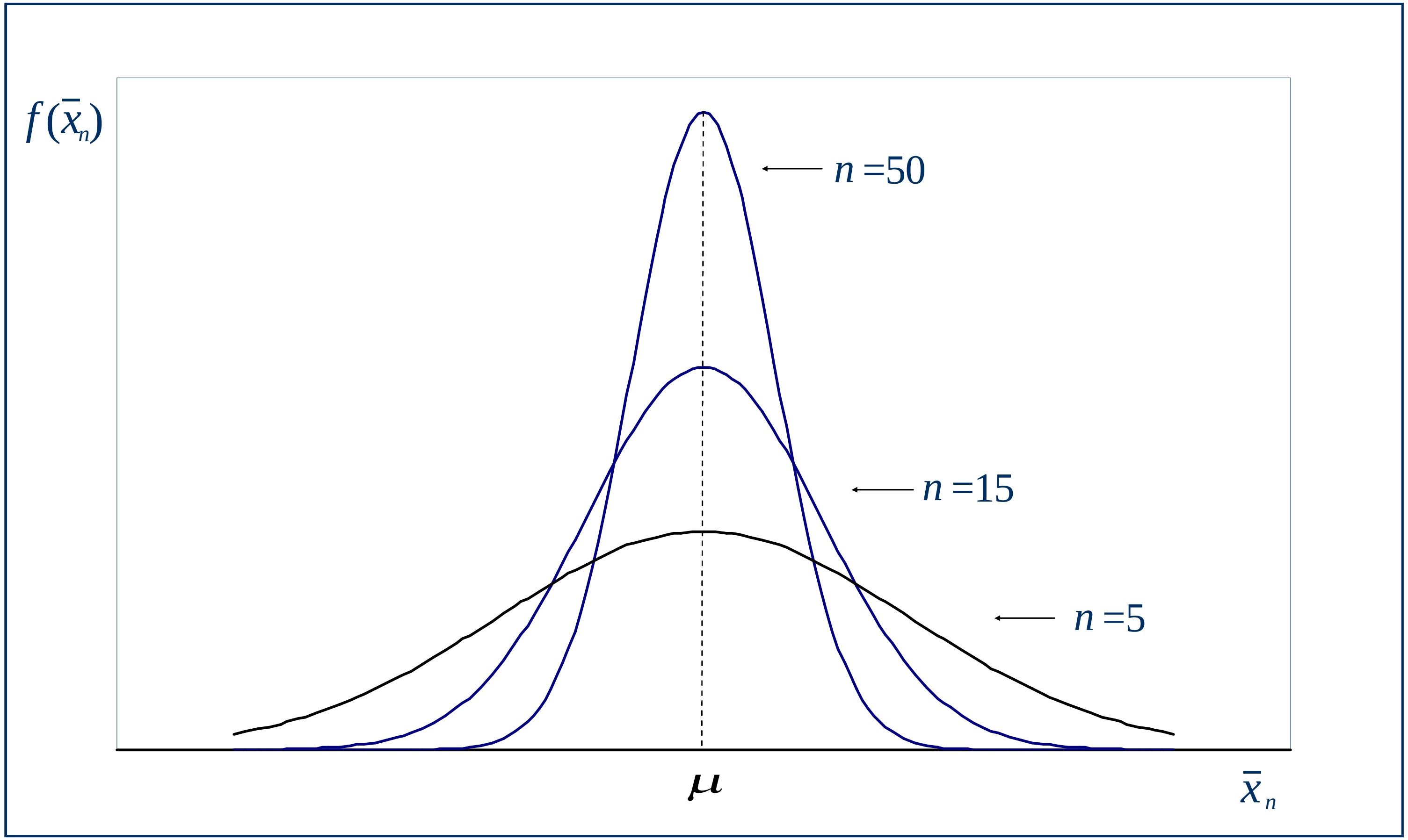
一致性估計式(強則)

設 $\hat{\theta}_n$ 為樣本數 n 之 θ 估計式，若 $P(\lim_{n\to\infty}\hat{\theta}_n=\theta)=1$，則$\hat{\theta}_n$為$\theta$之一致性估計式。

一致性估計式(弱則)

設 $\hat{\theta}_n$ 為樣本數 n 之 θ 估計式，若 $\lim_{n\to\infty}P(|\hat{\theta}_n-\theta|<\varepsilon)=1$，則$\hat{\theta}_n$為$\theta$之一致性估計式。其中$\varepsilon$為正的極小數值。

圖10.12 一致性估計式



定理

若 $\hat{\theta}_n$ 為不偏誤估計式或不偏漸近估計式,且當n趨於無窮大時, 其變異數會趨近於零, 即 $\lim_{n\to\infty} V(\hat{\theta}_n) = 0$,則 $\hat{\theta}_n$ 為 θ的一致性估計式.

一致性的定理

若 $\hat{\theta}_n$ 為θ的一致性估計式,則任意的連續函數 $g(\hat{\theta}_n)$亦為g(θ)的一致估計式

林惠玲 陳正倉著 雙葉書廊發行 2004

EX10.16 請證明 $\bar{X}_n$ 與 S_n^2 分別爲μ, σ^2的一致性估計式

解: 因 $\bar{X}$ 爲μ的不偏估計式, 又知

$$V(\bar{X}_n) = \frac{\sigma^2}{n} \quad 且 \quad \lim_{n\to\infty}\frac{\sigma^2}{n} = 0$$

又S^2爲 σ^2的不偏估計式, 又知 (當 X 爲常態分配時)

$$V(S_n^2) = \frac{2\sigma^4}{n-1} \quad 且 \quad \lim_{n\to\infty}\frac{2\sigma^2}{n-1} = 0$$

根據上述定理, $\bar{X}_n$ 與 S_n^2 分別爲μ, σ^2的一致性估計

10.2.6 充分性

❍ 充分性

充分性是指若一估計式$\hat{\theta}$在估計θ時充分利用樣本資料的訊息，則稱$\hat{\theta}$為θ的充分估計式。

10.3 判斷估計式的應注意事項

1. 不偏性爲一個平均的性質

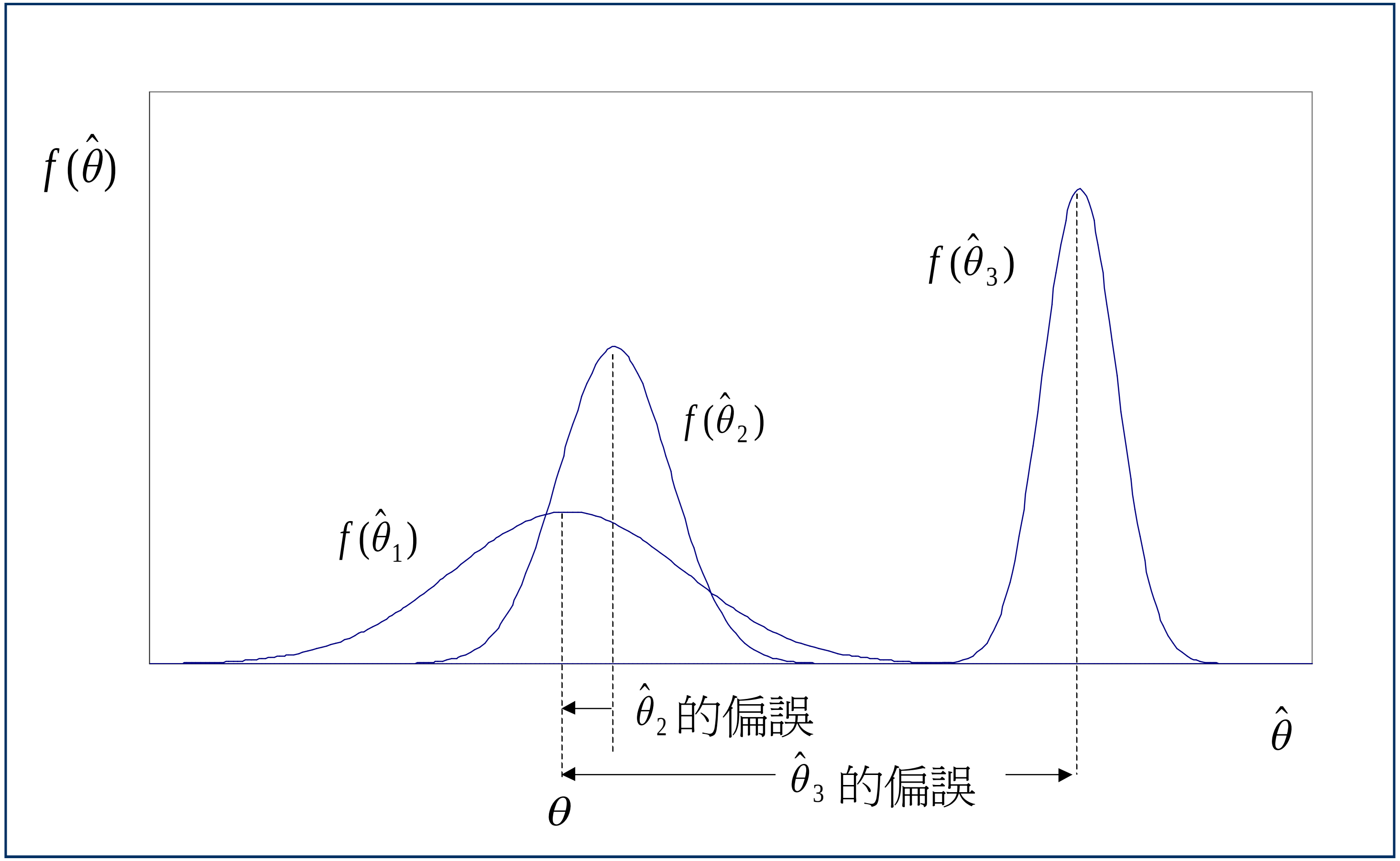
2. 一致性爲大樣本的性質

3. 最小變異不偏性具實用性

4. 評判標準以何者較爲重要尚無定論

林惠玲　陳正倉著 雙葉書廊發行 2004

圖10.13 三個估計式的比較



林惠玲 陳正倉著 雙葉書廊發行 2004

10.4 統計(點)估計的方法

較常用的方法爲: 最小平方法 (least square method)
動差法 (method of moment)
最大概似法 (maximum likelihood method)

10.4.1 最小平方法

概念: 找到一個估計式使估計值與觀察值的誤差平方和最小

Why? 1. 誤差有正負彼此會抵銷, 不易看出誤差的大小,因此取平方

2. 取平方有重視較大誤差的意義,因絕對數值越大平方會越大

林惠玲 陳正倉著 雙葉書廊發行 2004

10.4.2 動差法

設X為一隨機變數,其機率函數為f(x; θ); θ爲母體參數

- 母體r級原始動差

$$U_r = E(X^r)\text{，} r = 1,2,3\cdots$$

- 樣本r級原始動差

$$U_1' = \frac{\sum_{i=1}^{n} X_i^r}{n}\text{，} r = 1,2,3\cdots$$

- 動差法的性質

1. 動差估計式爲樣本動差的函數, 因樣本動差爲母體動差的一致性估計式, 因此可證明動差估計式必爲一致性估計式

2. 在一般情況下, 動差估計式的抽樣分配當 $n \to \infty$ 時會趨近常態分配

3. 動差估計式通常不具漸進有效性

4. 動差法通常必須知其母體分配 $f(x; \theta)$

林惠玲 陳正倉著 雙葉書廊發行 2004

10.4.3 最大概似法

❍ 概似函數

設$(X_1,\cdots,X_n)$為從已知母體$f(X;\theta)$簡單隨機抽出之一組隨機樣本，θ代表母體參數，未知；$(x_1,x_2,\cdots,x_n)$為其觀察值，則概似函數寫為：

$$
\begin{aligned}
L(x_1,x_2,\cdots,x_n;\theta) &= f(x_1;\theta)f(x_2;\theta)\ldots f(x_n;\theta) \\
&= \Pi_{i=1}^{n} f(x_i;\theta) \\
&= L(\theta)
\end{aligned}
$$

❍ 最大概似法

求使概似函數最大的估計式方法稱為最大概似法，即求$\hat{\theta}$使

$$\frac{\partial L(\theta)}{\partial\theta}=0 \quad 且 \quad \frac{\partial^2 L(\theta)}{\partial\theta^2}<0$$

林惠玲 陳正倉著 雙葉書廊發行 2004

表10.6　樣本出現的機率

母體參數（紅球數 θ ）	出現三紅球的機率 $L(\theta)$
3（3 紅 2 白）	$C_3^3 / C_3^5 = 1/10$
4（4 紅 1 白）	$C_3^4 / C_3^5 = 4/10$
5（5 紅）	$C_3^5 / C_3^5 = 1$

- 最大概似法的性質

1. MLE必為一致性估計式

2. MLE具漸近不偏性但不一定合於不偏性

3. MLE具漸進有效性

4.MLE通常必須知母體分配 $f(x;\theta)$, 求解複雜為其缺點