第四章

資料的整理與表現－統計測量數

學習目的

1. 了解資料中央趨勢的各種衡量指標，如算術平均數、中位數、眾數、加權平均數與幾何平均數等的衡量方法。

學習目的

1. 了解資料中央趨勢的各種衡量指標，如算術平均數、中位數、眾數、加權平均數與幾何平均數等的衡量方法。
2. 熟習各個中央趨勢衡量方法的特性、使用時機與優缺點。

學習目的

1. 了解資料中央趨勢的各種衡量指標，如算術平均數、中位數、眾數、加權平均數與幾何平均數等的衡量方法。
2. 熟習各個中央趨勢衡量方法的特性、使用時機與優缺點。
3. 了解資料分散程度的各種衡量指標，如全距、四分位距、變異數、標準差、變異係數的衡量方法。

學習目的

1. 了解資料中央趨勢的各種衡量指標，如算術平均數、中位數、眾數、加權平均數與幾何平均數等的衡量方法。
2. 熟習各個中央趨勢衡量方法的特性、使用時機與優缺點。
3. 了解資料分散程度的各種衡量指標，如全距、四分位距、變異數、標準差、變異係數的衡量方法。
4. 熟習各個分散程度衡量方法的特性、使用時機與優缺點。

學習目的

1. 了解資料中央趨勢的各種衡量指標，如算術平均數、中位數、眾數、加權平均數與幾何平均數等的衡量方法。
2. 熟習各個中央趨勢衡量方法的特性、使用時機與優缺點。
3. 了解資料分散程度的各種衡量指標，如全距、四分位距、變異數、標準差、變異係數的衡量方法。
4. 熟習各個分散程度衡量方法的特性、使用時機與優缺點。
5. 認識資料相對位置的各種衡量方法，如四分位數、十分位數、百分位數等的計算。認識與計算資料的偏度、峰度。

學習目的

1. 了解資料中央趨勢的各種衡量指標，如算術平均數、中位數、眾數、加權平均數與幾何平均數等的衡量方法。
2. 熟習各個中央趨勢衡量方法的特性、使用時機與優缺點。
3. 了解資料分散程度的各種衡量指標，如全距、四分位距、變異數、標準差、變異係數的衡量方法。
4. 熟習各個分散程度衡量方法的特性、使用時機與優缺點。
5. 認識資料相對位置的各種衡量方法，如四分位數、十分位數、百分位數等的計算。認識與計算資料的偏度、峰度。
6. 熟習使用EXCEL計算中央趨勢與分散度指標及其他位置之指標。

本章結構

- 資料的整理與表現－統計測量數
  - 未分組資料中央趨勢的衡量
    - 平均數
    - 中位數
    - 眾數
    - 中央趨勢統計測量數的比較與選擇
  - 未分組資料位置的衡量
    - 四分位數
    - 十分位數
    - 百分位數
  - 未分組資料分散度的衡量
    - 全距與四分位距
    - 平均絕對離差
    - 變異數
    - 標準差
    - 相對分散度
  - 柴比氏定理與經驗法則
  - Z 值法
  - 未分組資料偏度與峰度的衡量
    - 未分組資料偏度的衡量
    - 未分組資料峰度的衡量
  - 盒鬚圖分析法
  - 分組資料中央趨勢的衡量
    - 算術平均數
    - 中位數
    - 眾數
  - 分組資料分散度的衡量
    - 變異數
    - 標準差
  - 分組資料位置的衡量
    - 四分位數
    - 十分位數
    - 百分位數

未分組資料中央趨勢的衡量

- 算術平均數
  所有觀察值的總和除以觀察值的個數即為算術平均數。算術平均數在數線上代表資料的平衡點。

未分組資料中央趨勢的衡量

- 算術平均數
  所有觀察值的總和除以觀察值的個數即為算術平均數。算術平均數在數線上代表資料的平衡點。
  - 母體平均數

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

未分組資料中央趨勢的衡量

- 算術平均數

  所有觀察值的總和除以觀察值的個數即為算術平均數。算術平均數在數線上代表資料的平衡點。

  - 母體平均數

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

  - 樣本平均數

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

金融業與保險業的月薪

金融業	25	27	28	28	29	30	31	34
保險業	20	21	21	22	23	26	28	78

金融業與保險業的月薪

金融業	25	27	28	28	29	30	31	34
保險業	20	21	21	22	23	26	28	78

金融業：$\overline{X} = \dfrac{25 + 27 + \cdots + 34}{8} = 29$

金融業與保險業的月薪

金融業	25	27	28	28	29	30	31	34
保險業	20	21	21	22	23	26	28	78

金融業：$\overline{X} = \dfrac{25 + 27 + \cdots + 34}{8} = 29$

保險業：$\overline{X} = \dfrac{20 + 21 + \cdots + 78}{8} = 30$

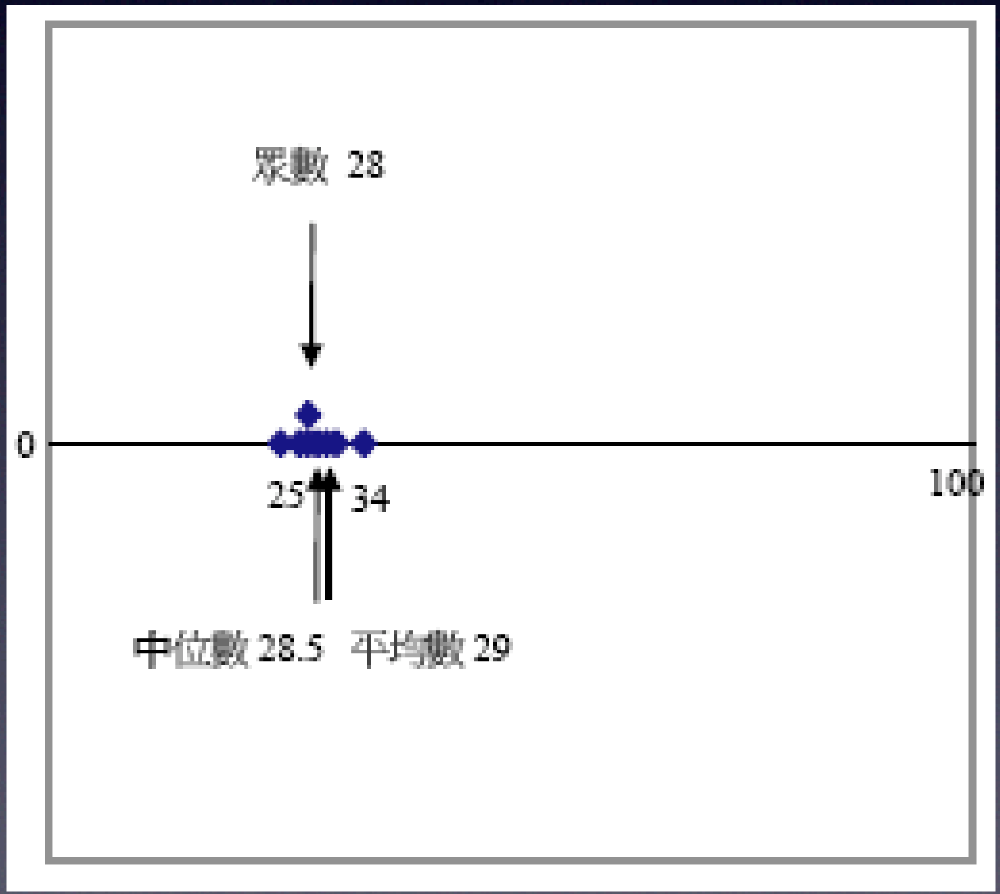
算術平均數的特質

1. 資料的平衡點。

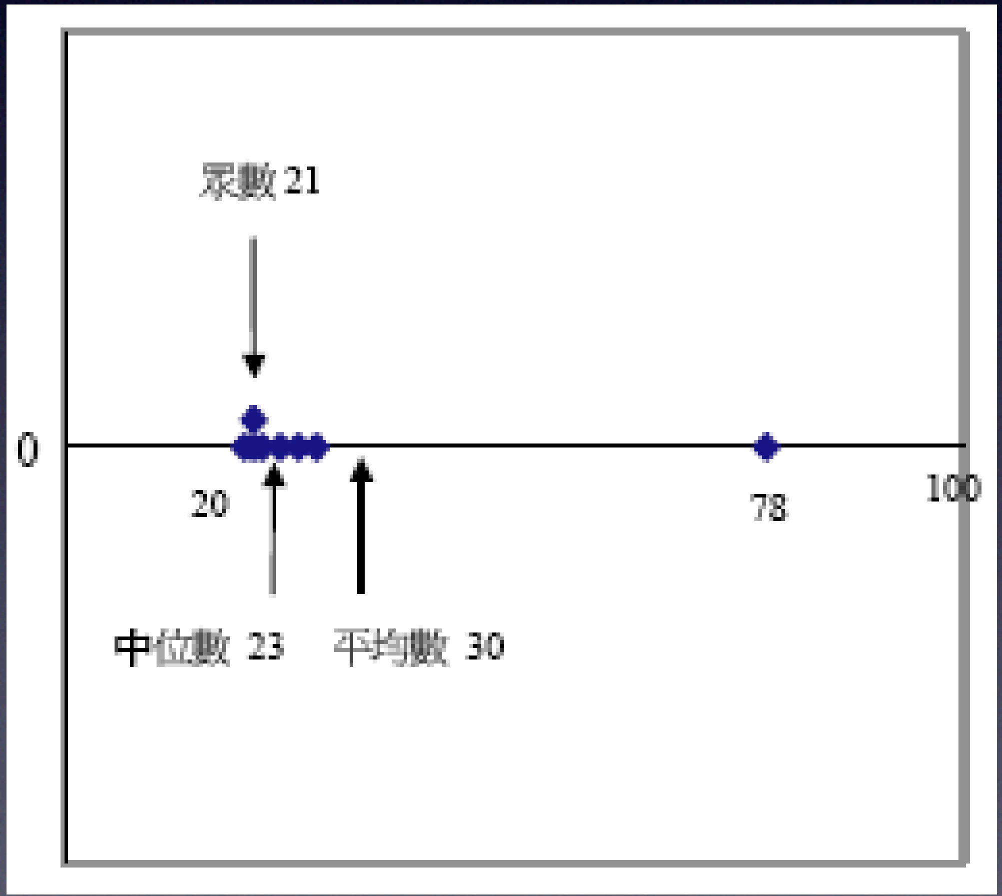
算術平均數的特質

1. 資料的平衡點。

金融業的平均月薪



保險業的平均月薪



算術平均數的特質

1. 資料的平衡點。
2. 各觀察值與平均數間的差的總和等於零。

算術平均數的特質

1. 資料的平衡點。
2. 各觀察值與平均數間的差的總和等於零。
3. 各觀察值與平均數之差的平方和最小。

算術平均數的特質

1. 資料的平衡點。
2. 各觀察值與平均數間的差的總和等於零。
3. 各觀察值與平均數之差的平方和最小。
4. 優點為考慮到每一個觀察值，缺點為易受極端值的影響。

算術平均數的特質

1. 資料的平衡點。
2. 各觀察值與平均數間的差的總和等於零。
3. 各觀察值與平均數之差的平方和最小。
4. 優點為考慮到每一個觀察值，缺點為易受極端值的影響。
5. 可進行代數演算。

算術平均數的特質

1. 資料的平衡點。
2. 各觀察值與平均數間的差的總和等於零。
3. 各觀察值與平均數之差的平方和最小。
4. 優點為考慮到每一個觀察值，缺點為易受極端值的影響。
5. 可進行代數演算。
6. 可對觀察值予以加權。

算術平均數的特質

1. 資料的平衡點。
2. 各觀察值與平均數間的差的總和等於零。
3. 各觀察值與平均數之差的平方和最小。
4. 優點為考慮到每一個觀察值，缺點為易受極端值的影響。
5. 可進行代數演算。
6. 可對觀察值予以加權。

母體加權算術平均數

$$\mu_W = \frac{\sum_{i=1}^{N} W_i x_i}{\sum_{i=1}^{N} W_i}$$

算術平均數的特質

1. 資料的平衡點。
2. 各觀察值與平均數間的差的總和等於零。
3. 各觀察值與平均數之差的平方和最小。
4. 優點為考慮到每一個觀察值，缺點為易受極端值的影響。
5. 可進行代數演算。
6. 可對觀察值予以加權。

樣本加權算術平均數

$$\overline{X_W} = \frac{\sum_{i=1}^{n} W_i x_i}{\sum_{i=1}^{n} W_i}$$

算術平均數的特質

科目	學分數	成績	加權成績	平均成績
國文	3	85	255	
英文	3	93	279	
歷史	2	76	152	
體育	1	87	87	
微積分	3	91	273	
經濟學原理	4	83	332	
會計學	3	92	276	
藝術欣賞	2	86	172	
合計	21		1826	86.95

未分組資料中央趨勢的衡量

- **幾何平均數**
  用來求等比數列的平均數，如百分比比率、指數等的平均數，以及求算某一期間平均成長率等。

未分組資料中央趨勢的衡量

- 幾何平均數
  用來求等比數列的平均數，如百分比比率、指數等的平均數，以及求算某一期間平均成長率等。
  - 母體的幾何平均數

$$G = \sqrt[N]{x_1 x_2 \cdots x_N} = \left(\prod_{i=1}^{N} x_i\right)^{\frac{1}{N}}$$

未分組資料中央趨勢的衡量

- 幾何平均數
  用來求等比數列的平均數，如百分比比率、指數等的平均數，以及求算某一期間平均成長率等。
  - 母體的幾何平均數

$$G = \sqrt[N]{x_1 x_2 \cdots x_N} = \left(\prod_{i=1}^{N} x_i\right)^{\frac{1}{N}}$$

  - 樣本的幾何平均數

$$g = \sqrt[n]{x_1 x_2 \cdots x_N} = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$$

台灣塑膠公司的股票價格

年度	台塑	變動比
87	46.5	
88	62.5	1.344
89	46.3	0.741
90	32.1	0.693
91	44.6	1.421
92	56	1.228

資料來源：台灣證券交易所。
註：價格的變動比為 P_t / P_{t-1}

台灣塑膠公司的股票價格

年度	台塑	變動比
87	46.5	
88	62.5	$\boxed{1.344}$
89	46.3	0.741
90	32.1	0.693
91	44.6	1.421
92	56	1.228

資料來源：台灣證券交易所。
註：價格的變動比為 $P_t / P_{t\text{-}1}$

$$\overline{g}_{\text{台塑}} = (1.344 \times 0.741 \times 0.693 \times 1.421 \times 1.228)^{\frac{1}{5}} = 1.038$$

幾何平均數的性質與應用

- 幾何平均數的性質
  - $\sqrt[n]{\prod_{i=1}^{n}(x_i/y_i)} = \sqrt[n]{\prod_{i=1}^{n}x_i} \Big/ \sqrt[n]{\prod_{i=1}^{n}y_i}$

幾何平均數的性質與應用

- 幾何平均數的性質
  - $\sqrt[n]{\prod_{i=1}^{n}(x_i/y_i)} = \sqrt[n]{\prod_{i=1}^{n}x_i} \Big/ \sqrt[n]{\prod_{i=1}^{n}y_i}$
  - 適合衡量等比數列的中央位置，但不易進行統計推論。

幾何平均數的性質與應用

- 幾何平均數的性質
  - $\sqrt[n]{\prod_{i=1}^{n}(x_i/y_i)} = \sqrt[n]{\prod_{i=1}^{n}x_i} \Big/ \sqrt[n]{\prod_{i=1}^{n}y_i}$
  - 適合衡量等比數列的中央位置，但不易進行統計推論。
- 幾何平均數的應用
  - 幾何平均數的投資報酬率

$$\overline{G} = [(1+R_1)(1+R_2)\cdots(1+R_n)]^{\frac{1}{n}} - 1$$

未分組資料中央趨勢的衡量

- **中位數的意義**
  中位數是位於依數值大小順序排列的觀察值中央的那一個數值。

未分組資料中央趨勢的衡量

- **中位數的意義**
  中位數是位於依數值大小順序排列的觀察值中央的那一個數值。

- **眾數的意義**
  眾數是指觀察值中，其出現次數最多的那一個數值。

未分組資料中央趨勢的衡量

- **中位數的意義**
  中位數是位於依數值大小順序排列的觀察值中央的那一個數值。

- **眾數的意義**
  眾數是指觀察值中，其出現次數最多的那一個數值。

台灣人民對兩岸關係的看法

看法	維持現狀以後再決定	永遠維持現狀	維持現狀以後統一	維持現狀以後獨立	儘快宣佈獨立	儘快統一
人數	358	231	111	108	63	45
百分比	39.05%	24.22%	12.12%	11.79%	6.88%	4.91%

資料來源：行政院陸委會91年12月委託e社會資訊管理公司調查。

中央趨勢統計測量數的比較

統計測量數	優點	缺點
算術平均數	1. 資料的重心。資料無極端值或偏態時，具代表性。 2. 適合代數演算。 3. 考慮所有觀察值，敏感度高。 4. 觀察值與平均數差平方和最小。 5. 適合統計推論的工作。	1. 若有極端值存在時，則不具代表性。 2. 資料如為偏態，則代表性較差。

中央趨勢統計測量數的比較

統計測量數	優點	缺點
算術平均數	1. 資料的重心。資料無極端值或偏態時，具代表性。 2. 適合代數演算。 3. 考慮所有觀察值，敏感度高。 4. 觀察值與平均數差平方和最小。 5. 適合統計推論的工作。	1. 若有極端值存在時，則不具代表性。 2. 資料如為偏態，則代表性較差。
幾何平均數	1. 適合等比資料。 2. 敏感度高。	1. 不適合一般資料。 2. 不適合作統計推論。

中央趨勢統計測量數的比較

統計測量數	優點	缺點
算術平均數	1. 資料的重心。資料無極端值或偏態時，具代表性。 2. 適合代數演算。 3. 考慮所有觀察值，敏感度高。 4. 觀察值與平均數差平方和最小。 5. 適合統計推論的工作。	1. 若有極端值存在時，則不具代表性。 2. 資料如為偏態，則代表性較差。
幾何平均數	1. 適合等比資料。 2. 敏感度高。	1. 不適合一般資料。 2. 不適合作統計推論。
中位數	1. 適用於有極端值的資料。 2. 適用於偏態資料。 3. 觀察值與中位數絕對差和最小。 4. 可做無母數統計推論。	1. 不適合代數演算。 2. 對觀察值敏感性低。 3. 不易進行母數統計推論。

中央趨勢統計測量數的比較

統計測量數	優點	缺點
算術平均數	1. 資料的重心。資料無極端值或偏態時，具代表性。 2. 適合代數演算。 3. 考慮所有觀察值，敏感度高。 4. 觀察值與平均數差平方和最小。 5. 適合統計推論的工作。	1. 若有極端值存在時，則不具代表性。 2. 資料如為偏態，則代表性較差。
幾何平均數	1. 適合等比資料。 2. 敏感度高。	1. 不適合一般資料。 2. 不適合作統計推論。
中位數	1. 適用於有極端值的資料。 2. 適用於偏態資料。 3. 觀察值與中位數絕對差和最小。 4. 可做無母數統計推論。	1. 不適合代數演算。 2. 對觀察值敏感性低。 3. 不易進行母數統計推論。
眾數	1. 適用於有極端值的資料。 2. 適用於偏態資料。 3. 適用於值的資料。	1. 可能不只一個或不存在。 2. 敏感性低。 3. 不能作統計推論。

士林與桃園地院訴訟案件審理日數

士林地院								桃園地院					
65	110	58	72	66	67	68	65	73	78	64	77	63	73
78	90	82	86					105	61	81	59		

士林與桃園地院訴訟案件審理日數

士林地院								桃園地院					
65	110	58	72	66	67	68	65	73	78	64	77	63	73
78	90	82	86					105	61	81	59		

地院訴訟案件審理日數的平均數中位數與眾數

士林地院		桃園地院		
平均數	75.58	平均數	73.4	
中間值	70	中間值	73	←中位數
眾數	65	眾數	73	

未分組資料位置的衡量（其他測量數）

- 四分位數
  四分位數是將順序資料分成四等分數值的分位數。

未分組資料位置的衡量（其他測量數）

- 四分位數
  四分位數是將順序資料分成四等分數值的分位數。
- 十分位數
  十分位數是將資料均分為十等份數值的分割數。

產業經濟學的學期成績

78	79	80	81	82	83	83	84	84
85	86	87	88	89	90	91	92	95

產業經濟學的學期成績

78	79	80	81	82	83	83	84	84
85	86	87	88	89	90	91	92	95

$$K_5 = \frac{n \cdot i}{10} = \frac{18 \times 5}{10} = 9$$

產業經濟學的學期成績

78	79	80	81	82	83	83	84	84
85	86	87	88	89	90	91	92	95

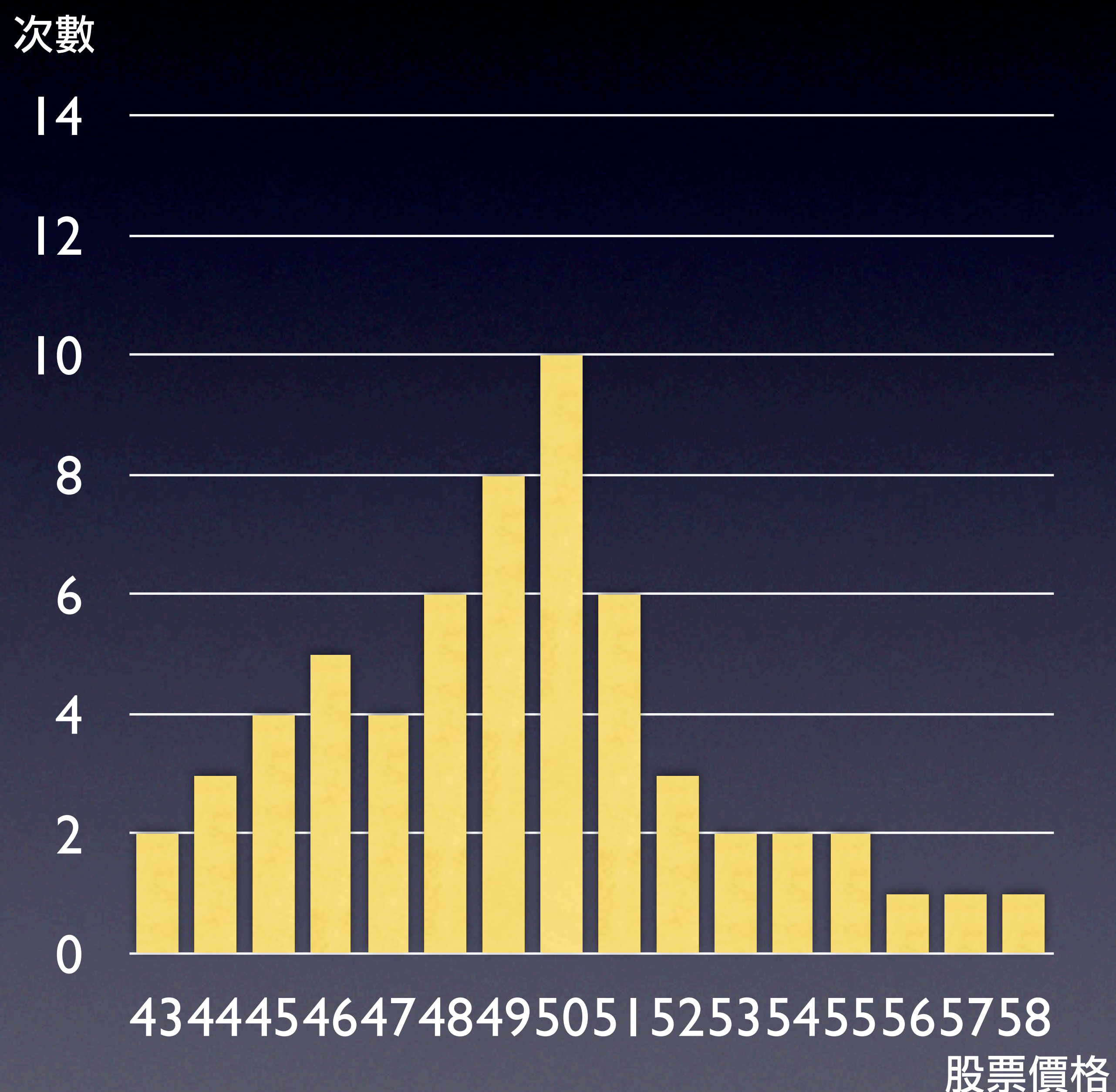
$$K_5 = \frac{n \cdot i}{10} = \frac{18 \times 5}{10} = 9$$

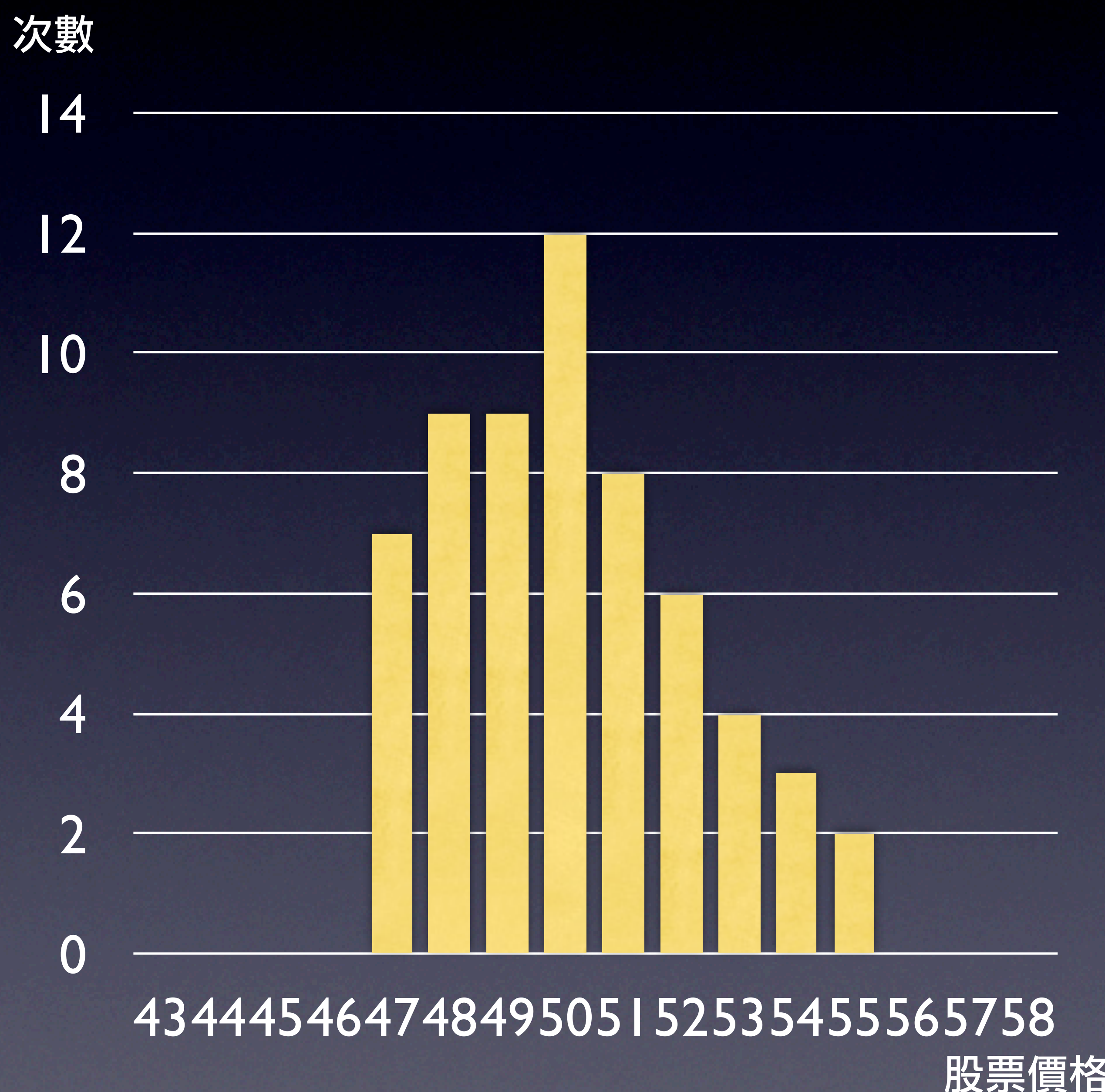
$$K_8 = \frac{n \cdot i}{10} = \frac{18 \times 8}{10} = 14.4$$

未分組資料位置的衡量（其他測量數）

- 四分位數
  四分位數是將順序資料分成四等分數值的分位數。
- 十分位數
  十分位數是將資料均分為十等份數值的分割數。
- 百分位數
  百分位數是將順序資料均分為一百等分數值的分割數。

未分組資料分散度的衡量





未分組資料分散度的衡量

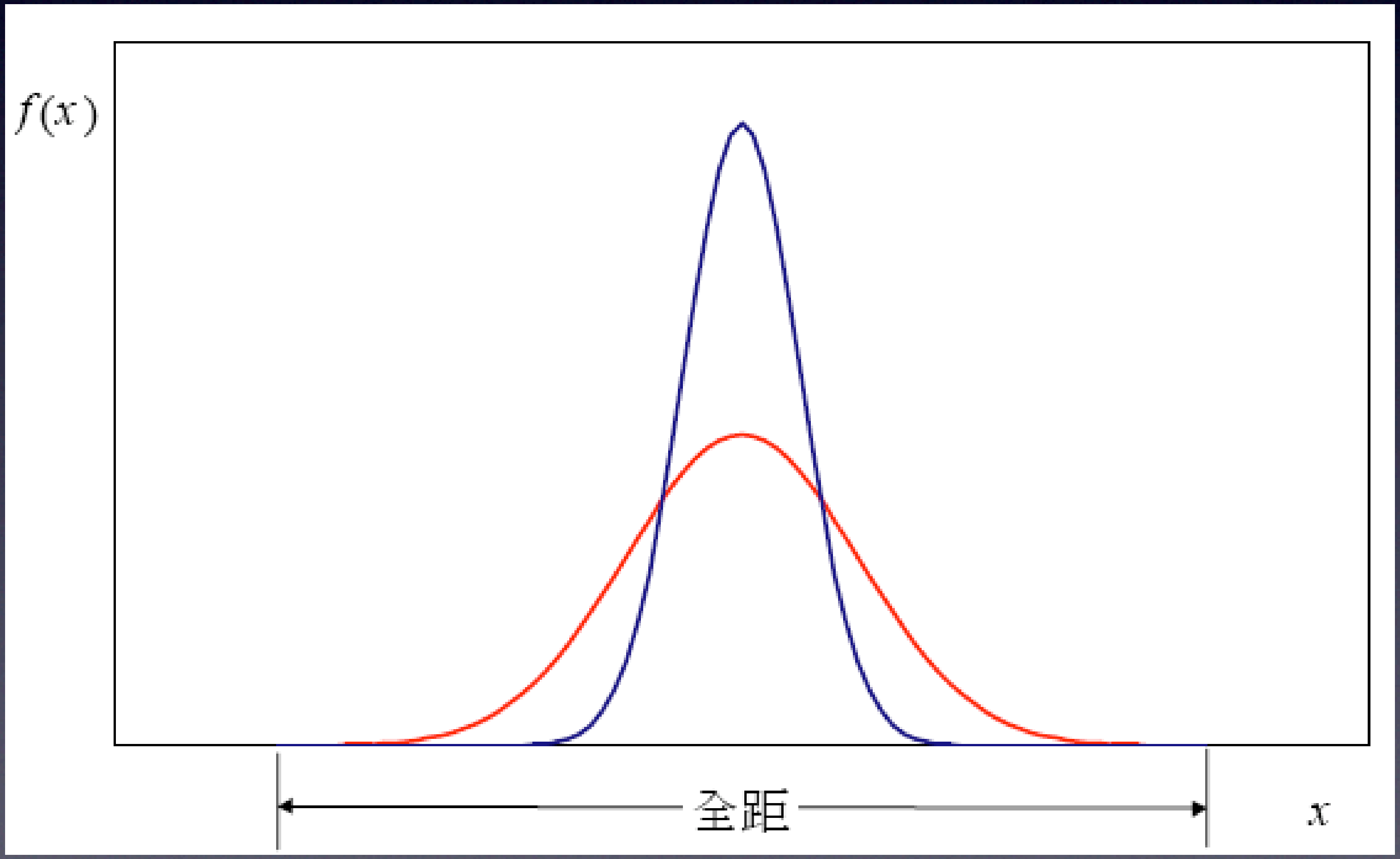
- 全距

  $R =$ 最大值 $-$ 最小值

未分組資料分散度的衡量

- 全距

  R = 最大值 − 最小值



未分組資料分散度的衡量

- 全距

  $R=$ 最大值 － 最小值

- 四分位距

  IQR = 第3四分位數 － 第1四分位數 $= Q_3 - Q_1$

未分組資料分散度的衡量

- 全距

  $R =$ 最大值 $-$ 最小值

- 四分位距

  IQR = 第3四分位數 $-$ 第1四分位數 $= Q_3 - Q_1$

- 平均絕對離差

  - 母體： $MAD = \frac{1}{N}\sum_{i=1}^{N}|x_i - \mu|$

未分組資料分散度的衡量

- 全距

  $R =$ 最大值 $-$ 最小值

- 四分位距

  IQR = 第3四分位數 $-$ 第1四分位數 $= Q_3 - Q_1$

- 平均絕對離差

  - 母體： $MAD = \frac{1}{N}\sum_{i=1}^{N}|x_i - \mu|$
  - 樣本： $mad = \frac{1}{n}\sum_{i=1}^{n}|x_i - \overline{X}|$

縱貫路與中山高的開車時間

縱貫路	40	42	43	44	46
中山高	21	36	42	55	61

縱貫路與中山高的開車時間

縱貫路	40	42	43	44	46
中山高	21	36	42	55	61

縱貫路開車時間的平均絕對離差

開車時間	$X-\overline{X}$	$\lvert X-\overline{X}\rvert$	$(X-\overline{X})^2$
40	-3	3	9
42	-1	1	1
43	0	0	0
44	1	1	1
46	3	3	9
合計	0	8	20

縱貫路與中山高的開車時間

縱貫路	40	42	43	44	46
中山高	21	36	42	55	61

縱貫路開車時間的平均絕對離差

開車時間	$X-\overline{X}$	$\lvert X-\overline{X}\rvert$	$(X-\overline{X})^2$
40	-3	3	9
42	-1	1	1
43	0	0	0
44	1	1	1
46	3	3	9
合計	0	8	20

$$mad=\frac{1}{n}\sum_{i=1}^{n}\lvert x_i-\overline{X}\rvert=\frac{8}{5}=1.6$$

未分組資料分散度的衡量

- 變異數
  - 母體變異數： $\sigma^2 = \frac{1}{N}\sum(x_i - \mu)^2$

未分組資料分散度的衡量

- 變異數
  - 母體變異數： $\sigma^2 = \frac{1}{N}\sum(x_i - \mu)^2$
  - 樣本變異數： $S^2 = \frac{1}{n-1}\sum(x_i - \overline{X})^2$

未分組資料分散度的衡量

- 變異數
  - 母體變異數： $\sigma^2 = \frac{1}{N}\sum(x_i - \mu)^2$
  - 樣本變異數： $S^2 = \frac{1}{n-1}\sum(x_i - \overline{X})^2$
- 標準差
  - 母體標準差： $\sigma = \sqrt{\sigma^2}$

未分組資料分散度的衡量

- 變異數
  - 母體變異數： $\sigma^2 = \frac{1}{N}\sum(x_i - \mu)^2$
  - 樣本變異數： $S^2 = \frac{1}{n-1}\sum(x_i - \overline{X})^2$
- 標準差
  - 母體標準差： $\sigma = \sqrt{\sigma^2}$
  - 樣本標準差： $S = \sqrt{S^2}$

未分組資料分散度的衡量

- **變異數**
  - 母體變異數： $\sigma^2 = \frac{1}{N}\sum(x_i - \mu)^2$
  - 樣本變異數： $S^2 = \frac{1}{n-1}\sum(x_i - \overline{X})^2$
- **標準差**
  - 母體標準差： $\sigma = \sqrt{\sigma^2}$
  - 樣本標準差： $S = \sqrt{S^2}$

縱貫路與中山高開車時間的比較

縱貫公路		中山高速	
平均數	43	平均數	43
中間值	43	中間值	42
眾數	N/A	眾數	N/A
標準差	2.24	標準差	15.83
變異數	5	變異數	250.5

未分組資料分散度的衡量

- 相對分散度
  - 變異係數： $CV = \dfrac{\text{標準差}}{\text{平均數}}$

未分組資料分散度的衡量

- 相對分散度
  - 變異係數： $CV = \dfrac{\text{標準差}}{\text{平均數}}$

母體資料：

$$CV = \frac{\sigma}{\mu}$$

樣本資料：

$$CV = \frac{S}{\overline{X}}$$

未分組資料分散度的衡量

- 相對分散度
  - 變異係數： $CV = \dfrac{\text{標準差}}{\text{平均數}}$

母體資料：

$$CV = \frac{\sigma}{\mu}$$

樣本資料：

$$CV = \frac{S}{\overline{X}}$$

兩種基金的平均數與標準差

基金類別	平均數 (%)	標準差 (%)	基金個數
跨國投資全球型	10.28	6.03	20
開放式價值類	5.09	3.71	5

資料來源：台灣經濟新報

未分組資料分散度的衡量

- 相對分散度
  - 變異係數： $CV = \dfrac{\text{標準差}}{\text{平均數}}$

母體資料：

$$CV = \frac{\sigma}{\mu}$$

樣本資料：

$$CV = \frac{S}{\overline{X}}$$

兩種基金的平均數與標準差

基金類別	平均數 (%)	標準差 (%)	基金個數
跨國投資全球型	10.28	6.03	20
開放式價值類	5.09	3.71	5

資料來源：台灣經濟新報

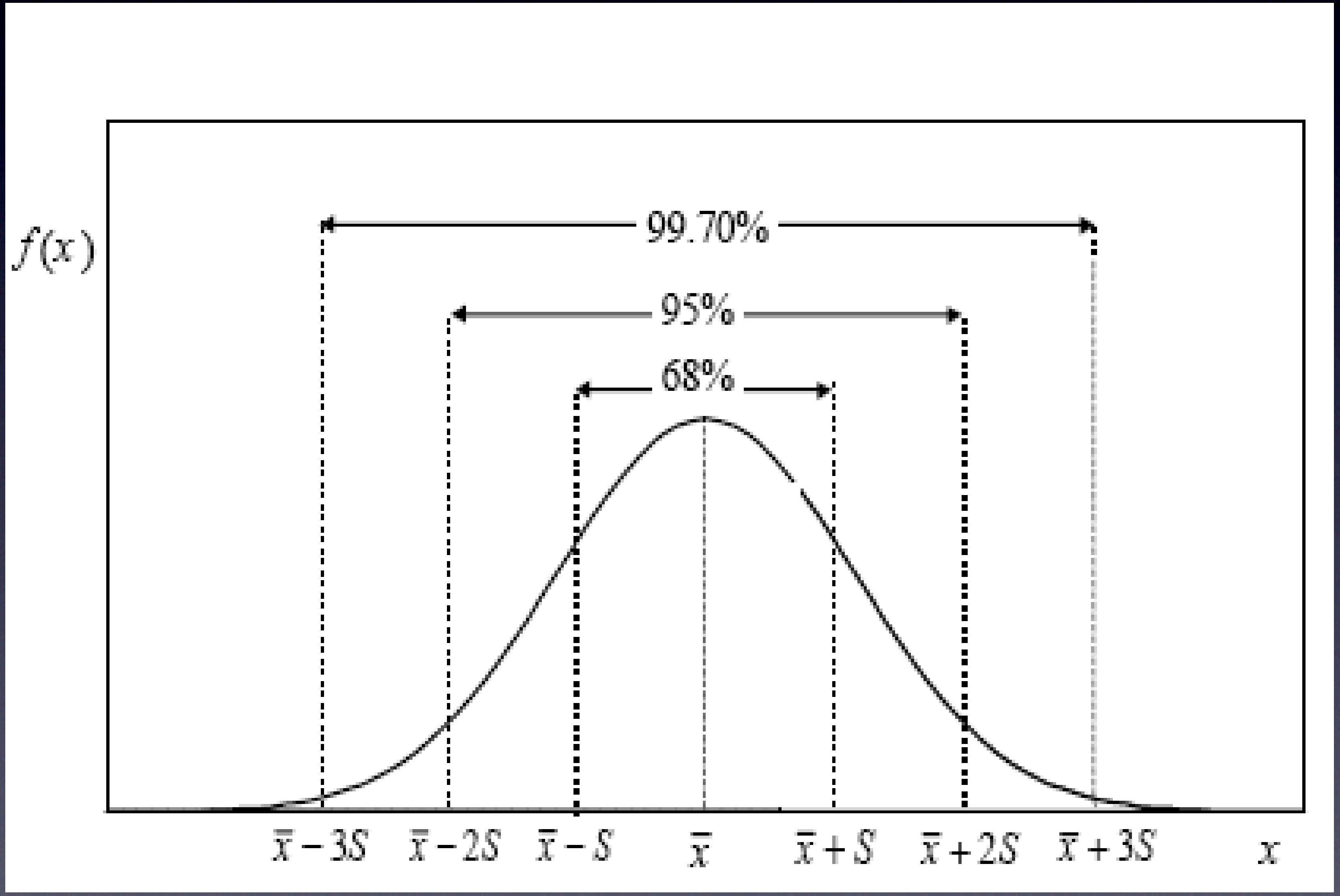
$$CV_1 = \frac{6.03}{10.28} = 0.59$$

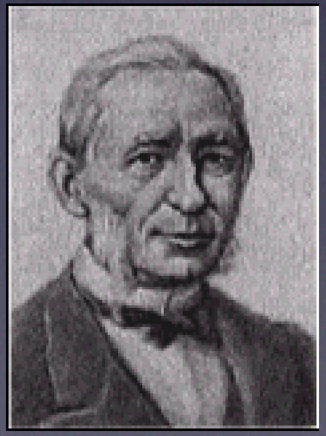
柴比氏定理與經驗法則

- 柴比氏定理
  不論資料為何種分配，至少有 $(1-1/k^2)$ 的資料落在距離平均數 k 個標準差的範圍內。k 為大於 1 的任意數，即 $k>1$。

柴比氏定理與經驗法則

經驗法則





P. L. Chebyshev

Z 值

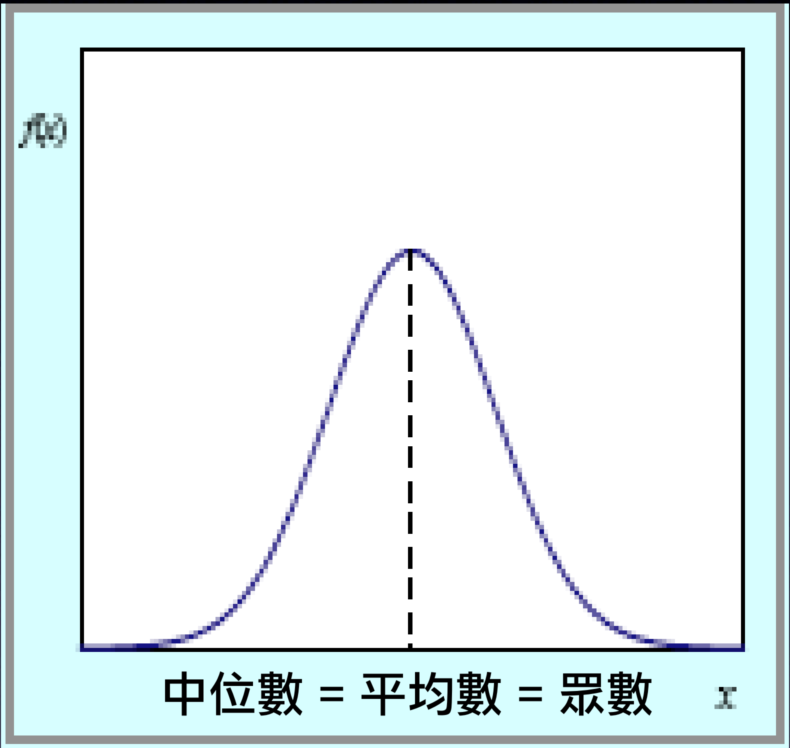
- 樣本 x 值的 Z 值：$\dfrac{x - \overline{x}}{s}$

Z 值

- 樣本 x 值的 Z 值： $\dfrac{x-\overline{x}}{s}$
- 母體 X 值的 Z 值： $\dfrac{X-\mu}{\sigma}$

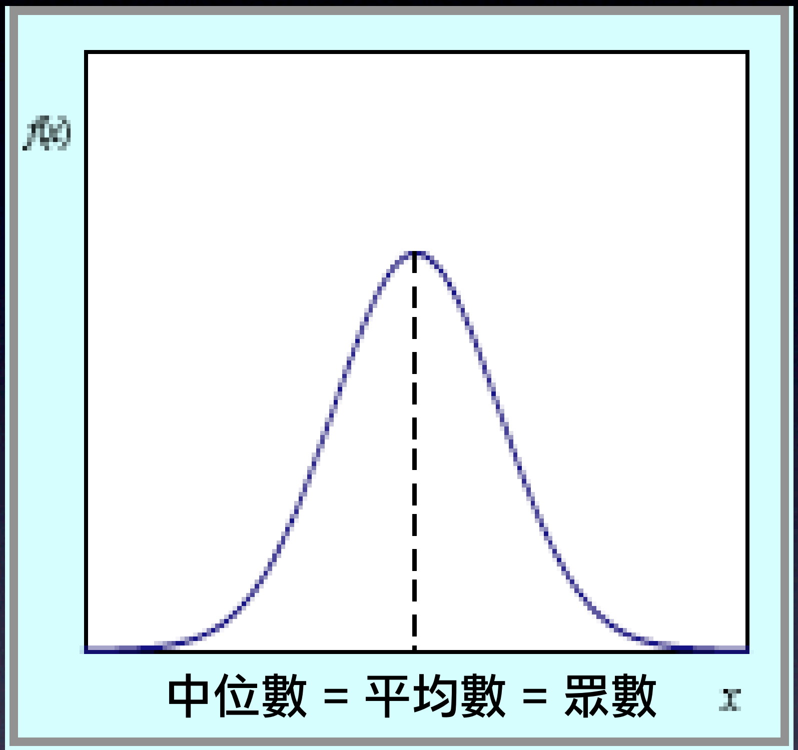
未分組資料偏度的測量

對稱分配

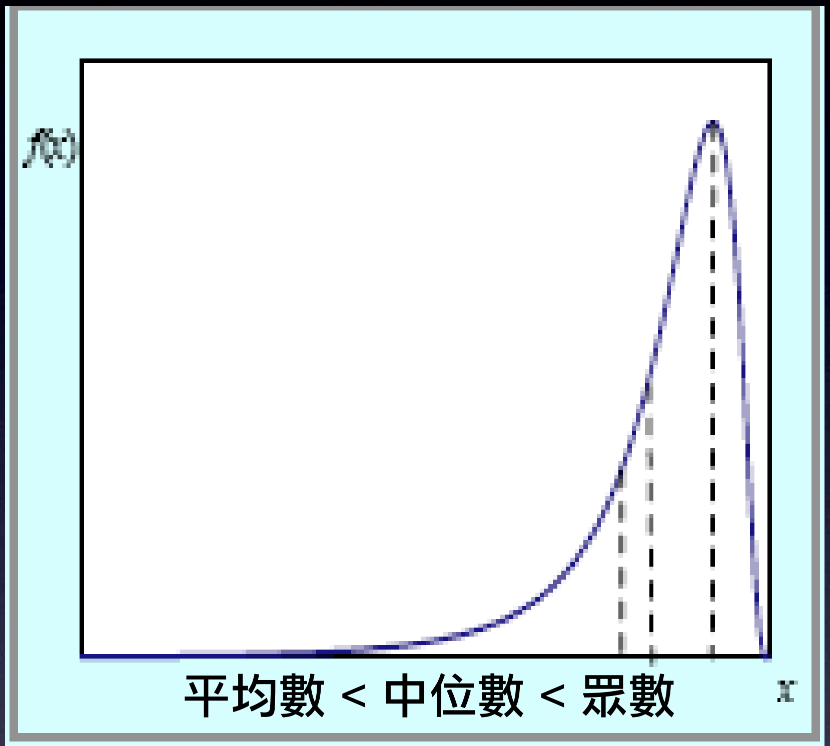


未分組資料偏度的測量

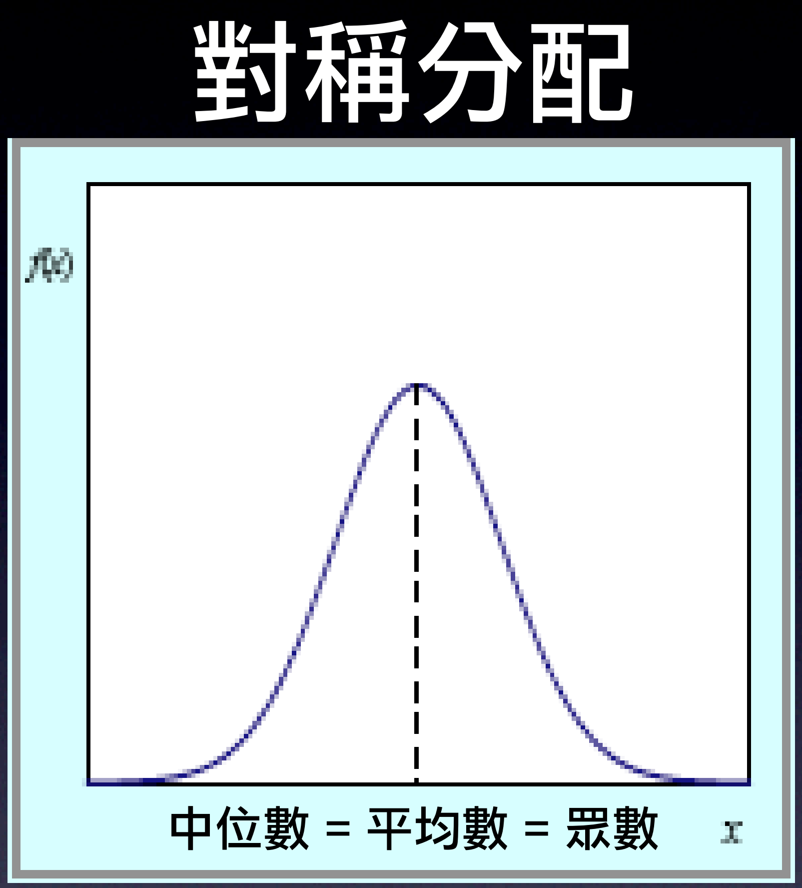
對稱分配

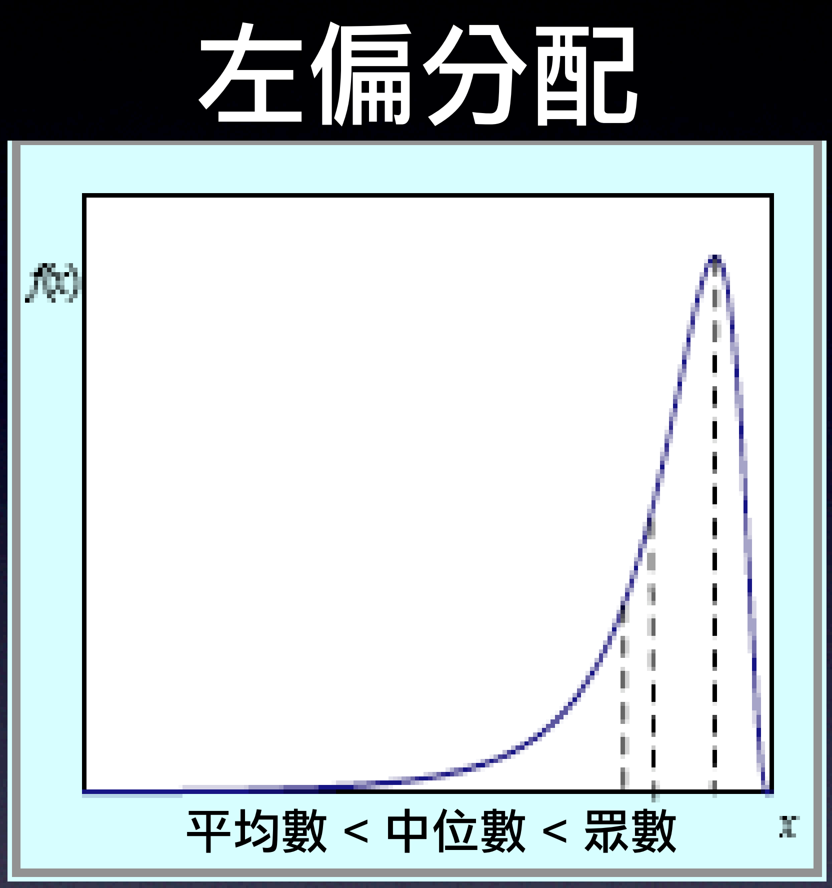


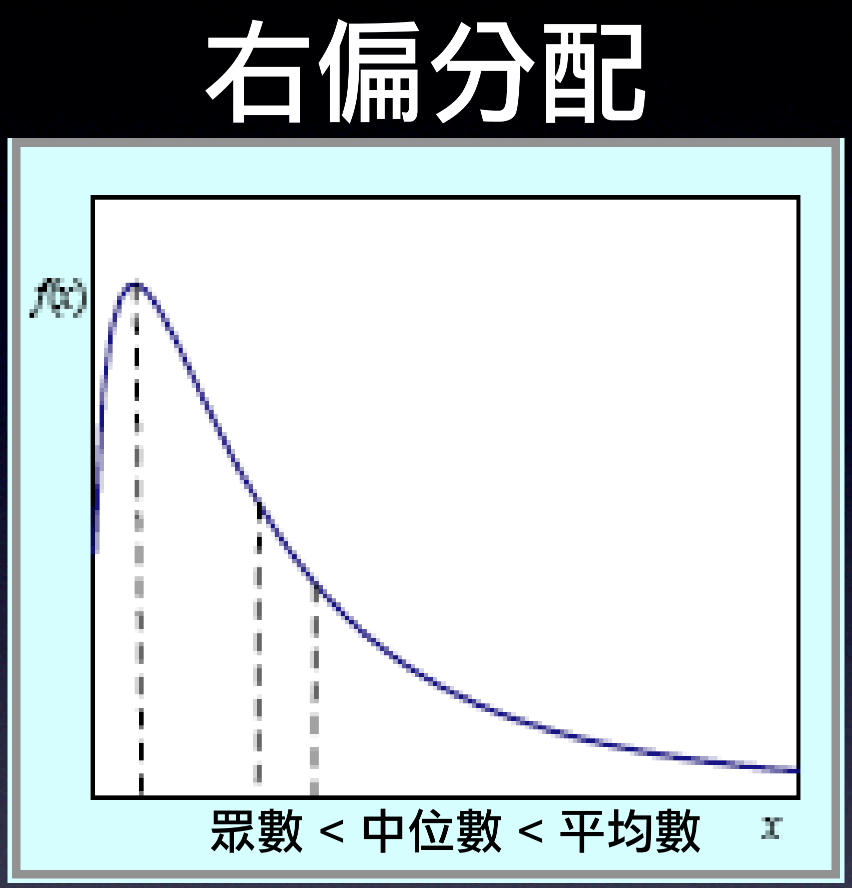
左偏分配



未分組資料偏度的測量

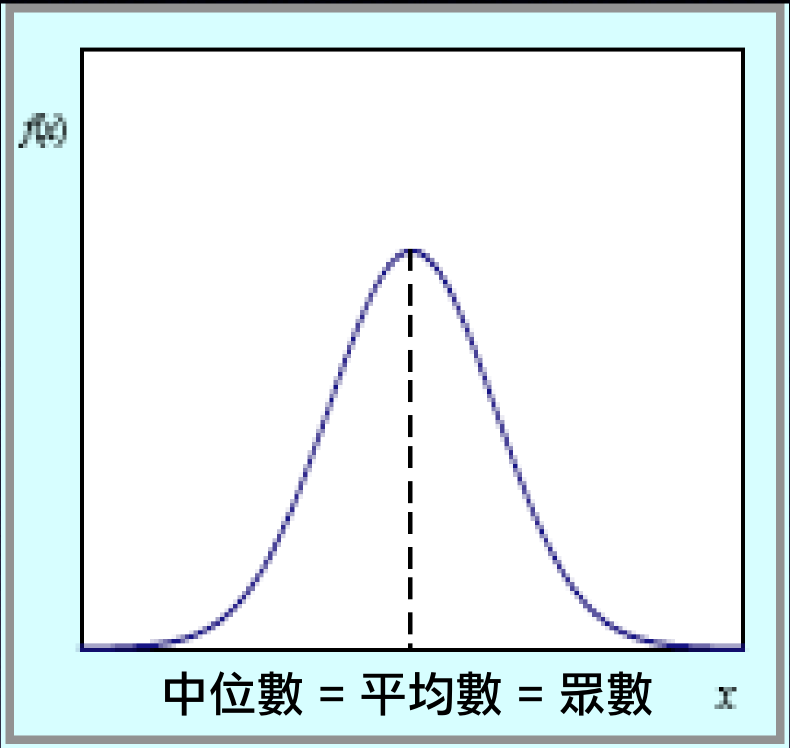




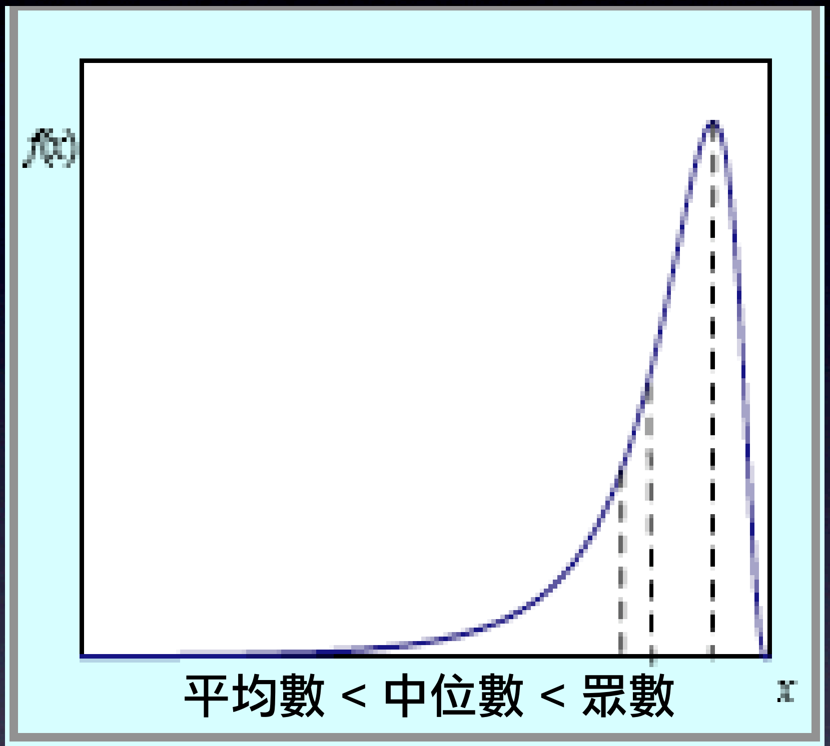


末分組資料偏度的測量

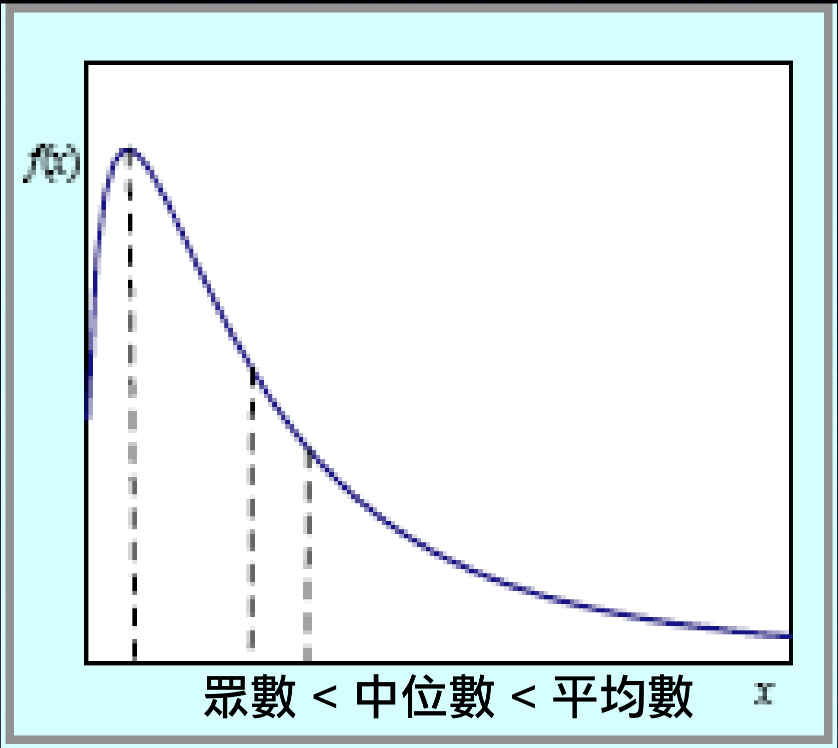
對稱分配



左偏分配



右偏分配



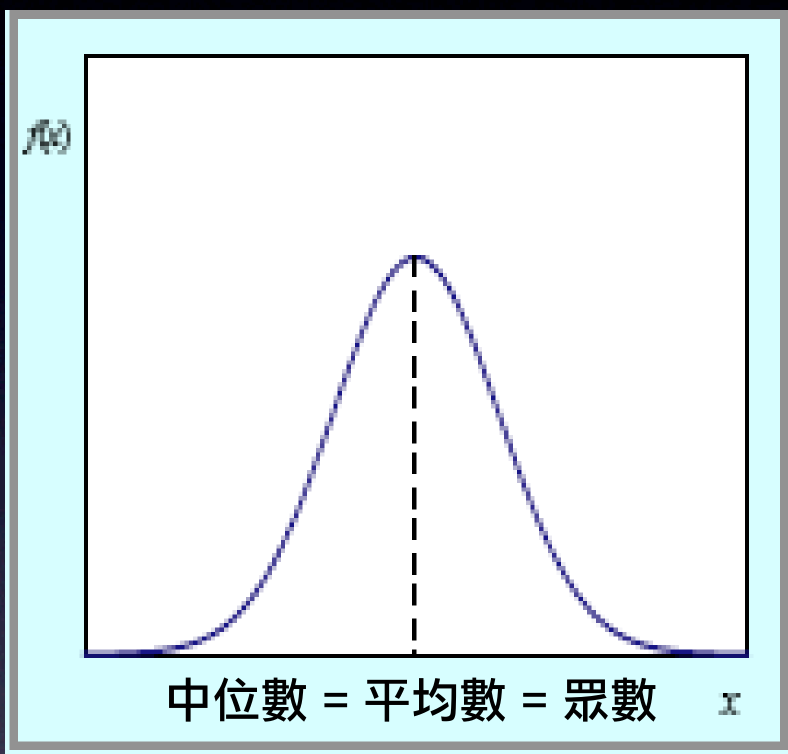
- 皮爾生偏態係數

  母體：

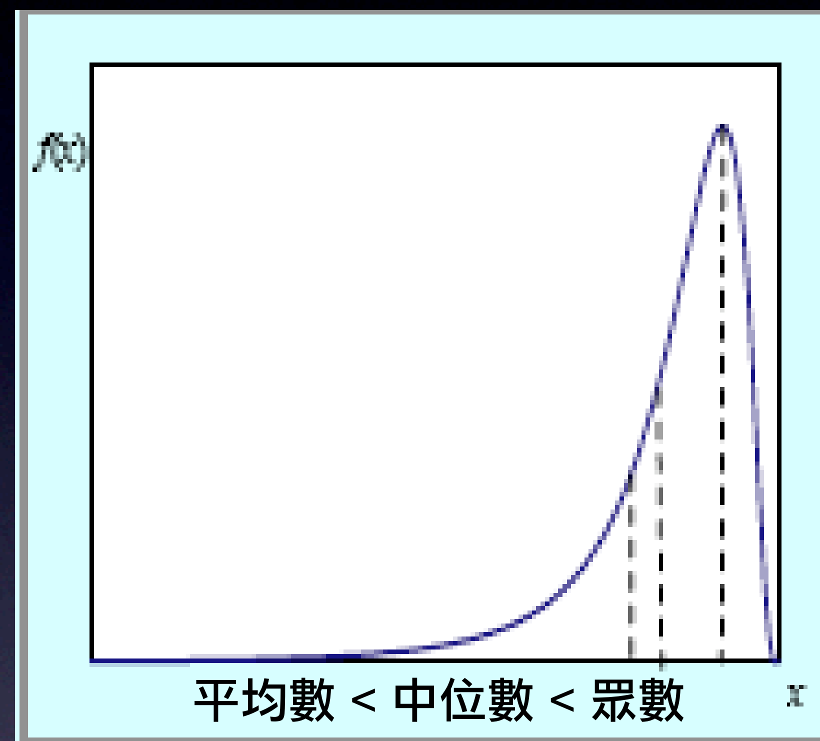
$$SK_p = \frac{3(\mu - M_e)}{\sigma}$$

未分組資料偏度的測量

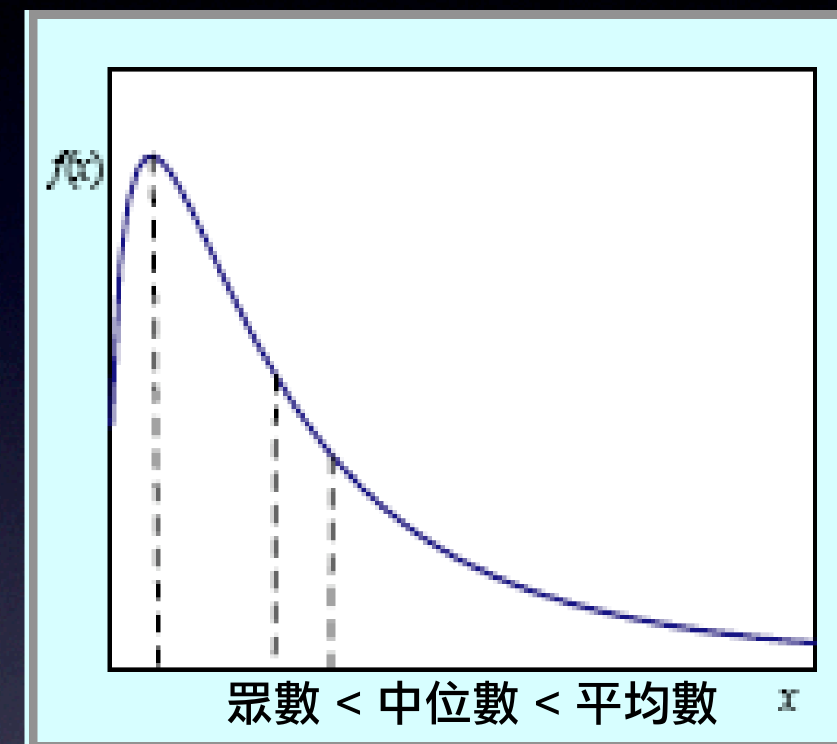
對稱分配



左偏分配



右偏分配



- 皮爾生偏態係數

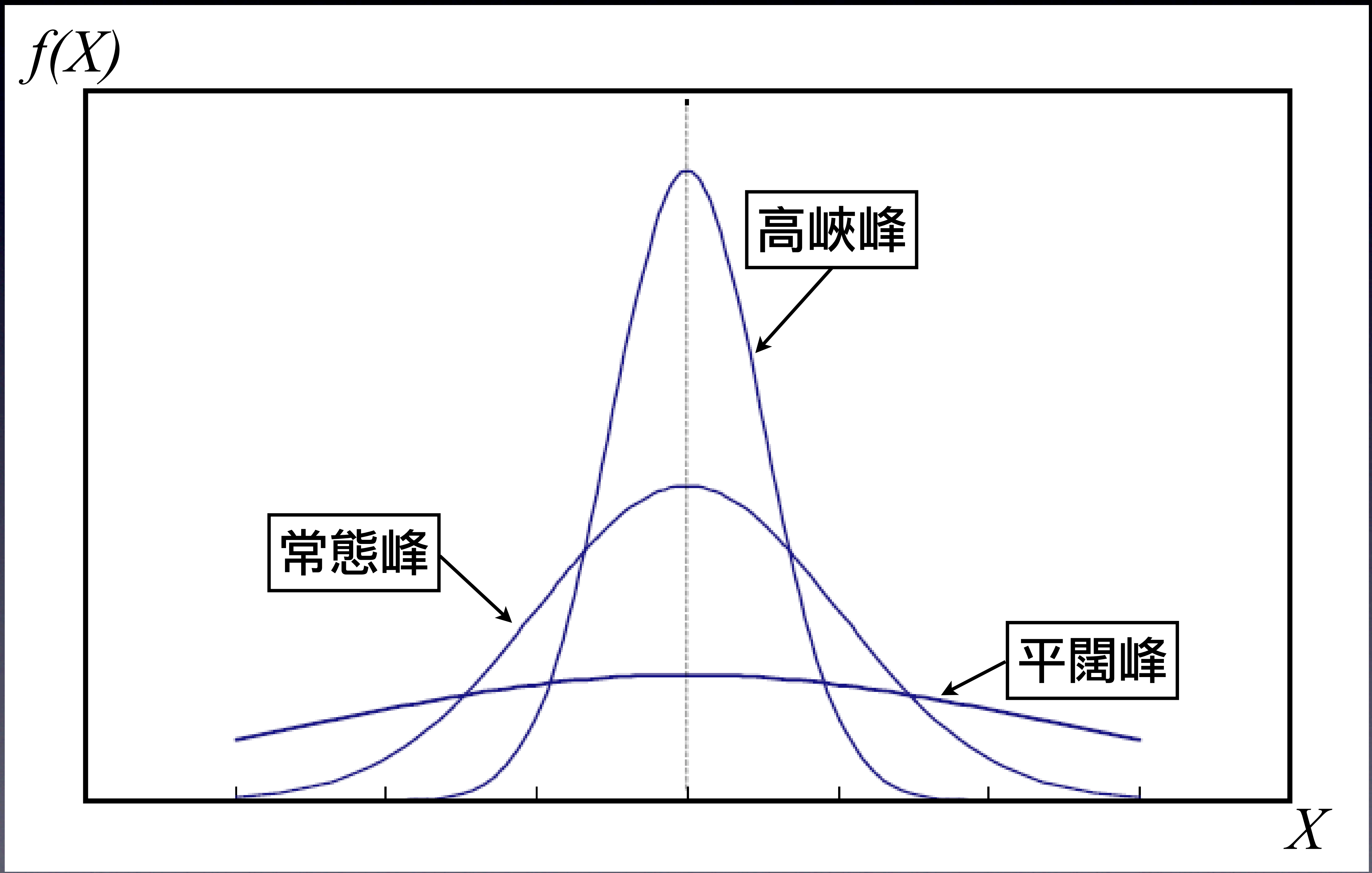
母體：

$$SK_p = \frac{3(\mu - M_e)}{\sigma}$$

樣本：

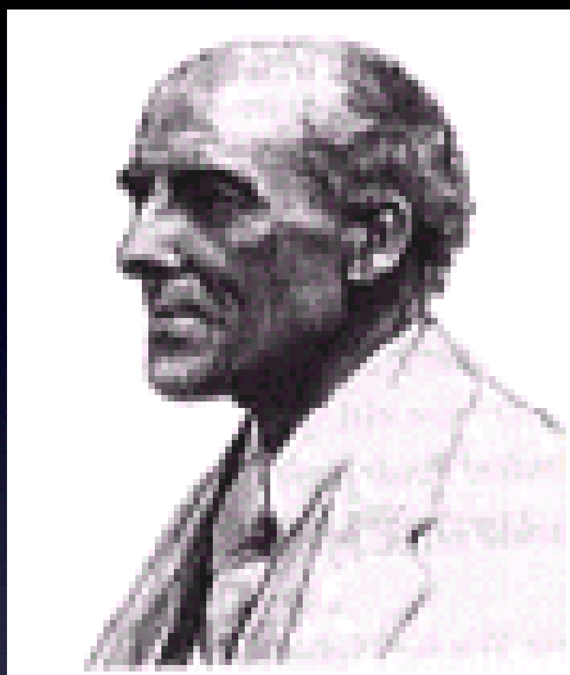
$$SK_p = \frac{3(\overline{X} - m_e)}{S}$$

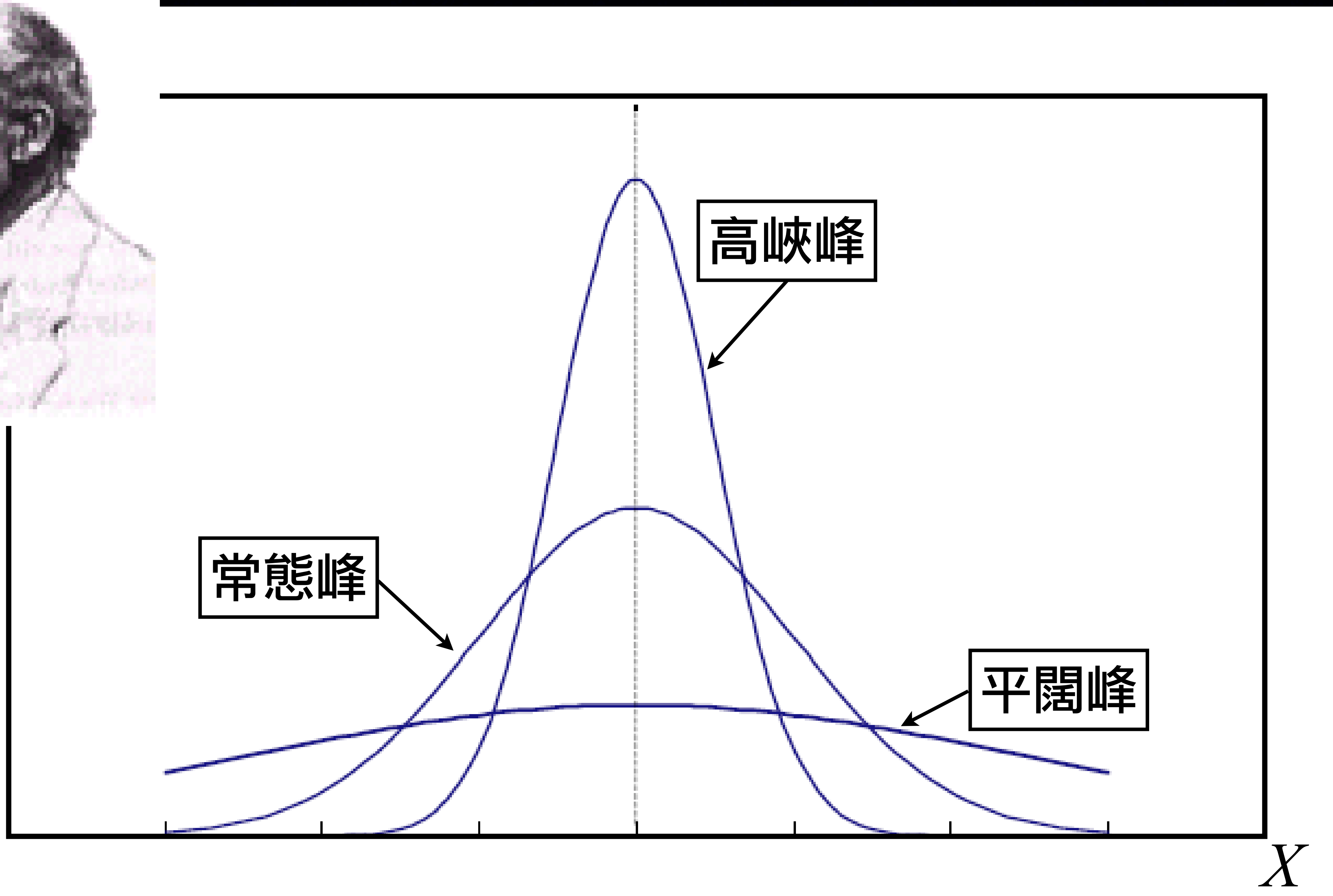
未分組資料峰度的測量



未分組資料峰度的測量

現代統計學的奠基者
Karl Pearson





應用統計學 林惠玲 陳正倉著 雙葉書廊發行 2006

盒鬚圖分析法

板橋及北投營業處的銷售業績

板橋營業處	32	52	64	64	70	76	82	280
北投營業處	61	72	81	81	95	97	101	124

盒鬚圖分析法

板橋及北投營業處的銷售業績

板橋營業處	32	52	64	64	70	76	82	280
北投營業處	61	72	81	81	95	97	101	124

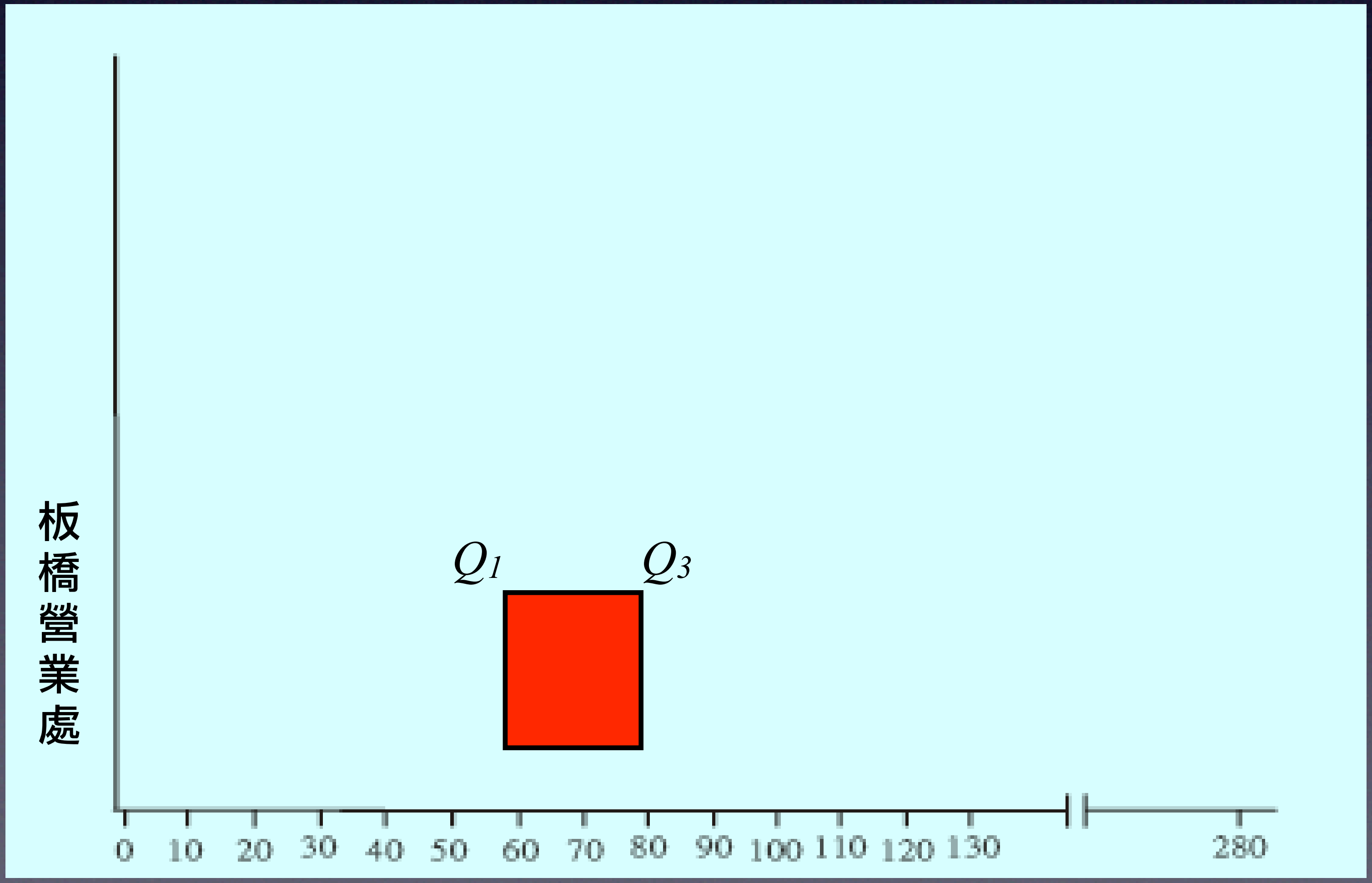
板橋：$min = 32$，$Q_1 = 58$，$M_e = 67$，$Q_3 = 79$，$max = 280$

盒鬚圖分析法

板橋及北投營業處的銷售業績

板橋營業處	32	52	64	64	70	76	82	280
北投營業處	61	72	81	81	95	97	101	124

板橋：$min = 32$，$Q_1 = 58$，$M_e = 67$，$Q_3 = 79$，$max = 280$

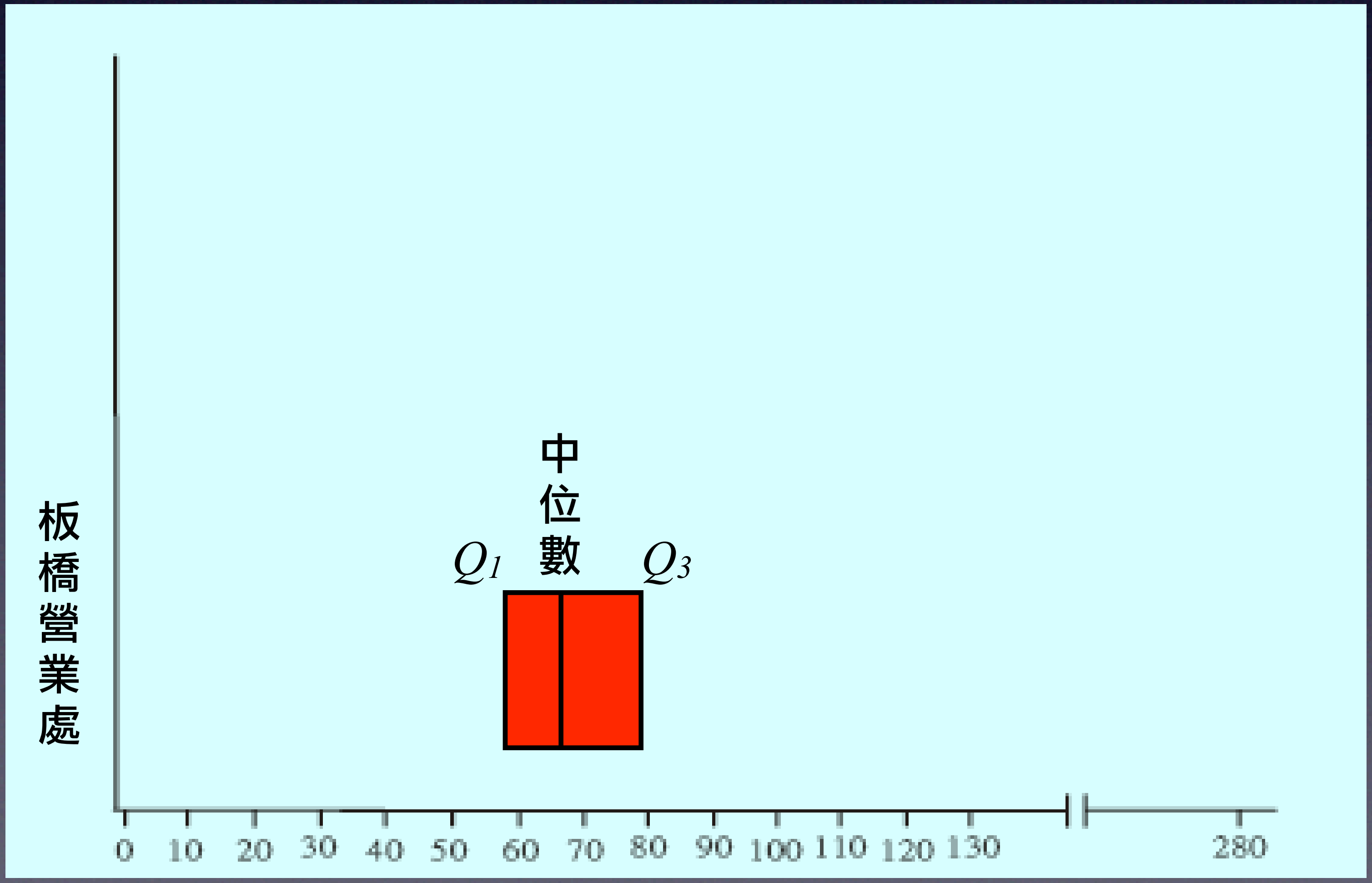


盒鬚圖分析法

板橋及北投營業處的銷售業績

板橋營業處	32	52	64	64	70	76	82	280
北投營業處	61	72	81	81	95	97	101	124

板橋：$min = 32$，$Q_1 = 58$，$M_e = 67$，$Q_3 = 79$，$max = 280$

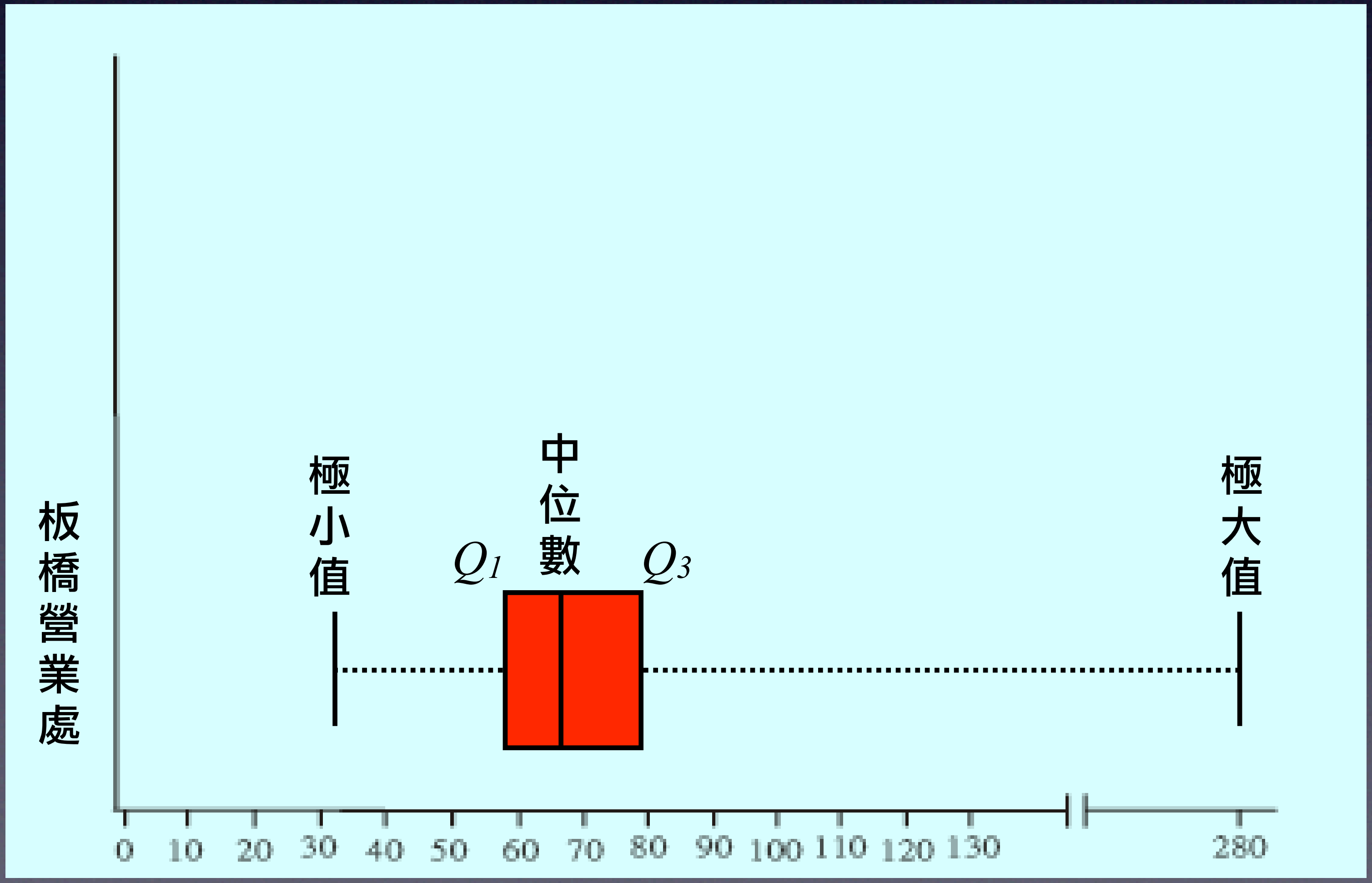


盒鬚圖分析法

板橋及北投營業處的銷售業績

板橋營業處	32	52	64	64	70	76	82	280
北投營業處	61	72	81	81	95	97	101	124

板橋：$min = 32$，$Q_1 = 58$，$M_e = 67$，$Q_3 = 79$，$max = 280$

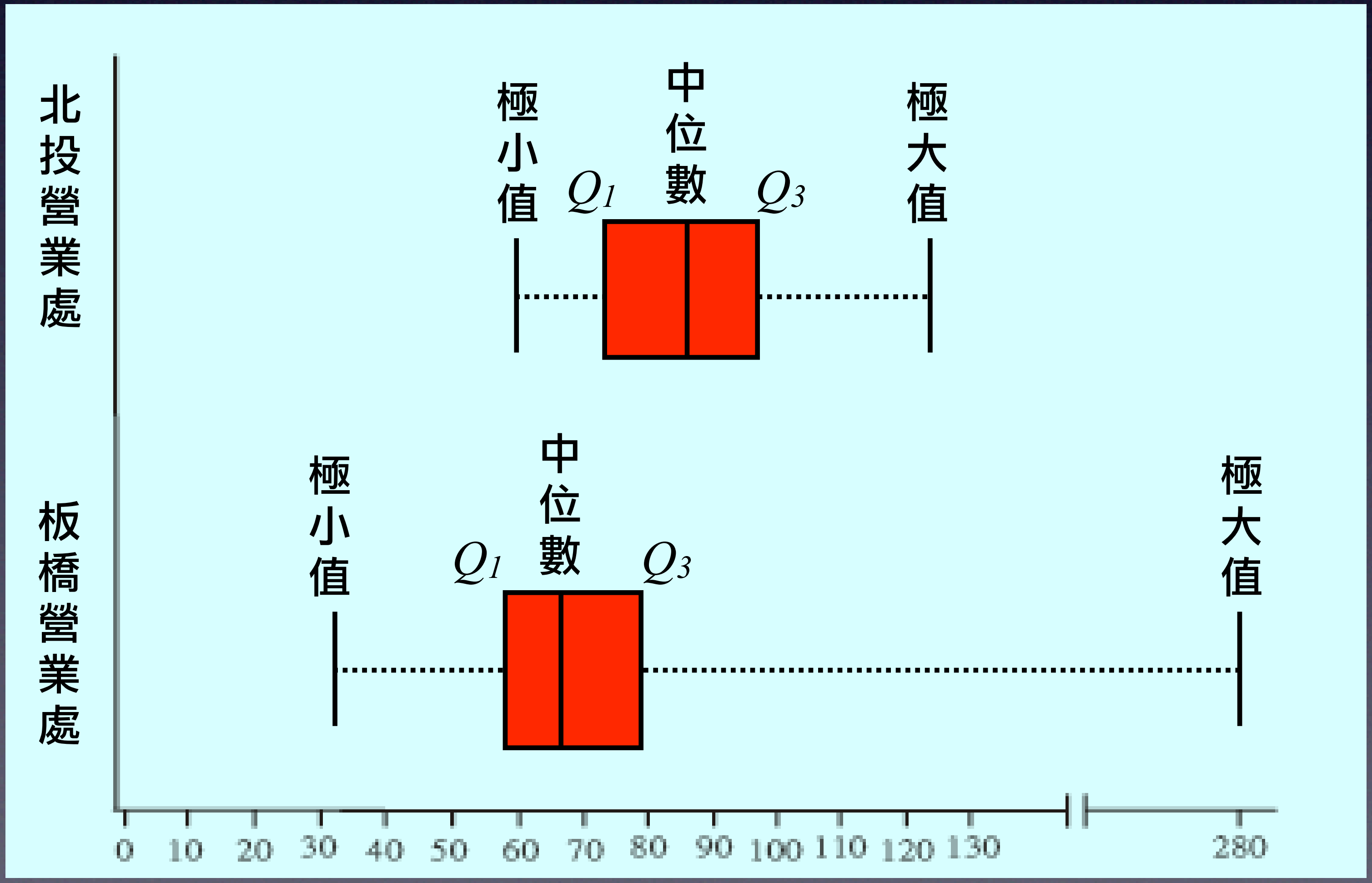


盒鬚圖分析法

板橋及北投營業處的銷售業績

板橋營業處	32	52	64	64	70	76	82	280
北投營業處	61	72	81	81	95	97	101	124

板橋：$min = 32$，$Q_1 = 58$，$M_e = 67$，$Q_3 = 79$，$max = 280$



應用統計學 林惠玲 陳正倉著 雙葉書廊發行 2006

分組資料中央趨勢的衡量

- 算術平均數

母體均數：$\mu = \dfrac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$　　樣本均數：$\overline{X} = \dfrac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$

股票型基金報酬率的次數分配表

組號	組限	組距	組中點	f_i 次數	$f_i x_i$	累加次數
1	$-45\% \le x < -30\%$	15%	-37.5%	1	-37.5	1
2	$-30\% \le x < -15\%$	15%	-22.5%	5	-112.5	6
3	$-15\% \le x < 0\%$	15%	-7.5%	18	-135	24
4	$0\% \le x < 15\%$	15%	7.5%	21	157.5	45
5	$15\% \le x < 30\%$	15%	22.5%	19	427.5	64
6	$30\% \le x < 45\%$	15%	37.5%	12	450	76
7	$45\% \le x < 60\%$	15%	52.5%	4	210	80
8	$60\% \le x < 75\%$	15%	67.5%	4	270	84
				$\Sigma f_i = 84$	1230	

股票型基金報酬率的次數分配表

組號	組限	組距	組中點	f_i 次數	$f_i x_i$	累加次數
1	$-45\% \leq x < -30\%$	15%	-37.5%	1	-37.5	1
2	$-30\% \leq x < -15\%$	15%	-22.5%	5	-112.5	6
3	$-15\% \leq x < 0\%$	15%	-7.5%	18	-135	24
4	$0\% \leq x < 15\%$	15%	7.5%	21	157.5	45
5	$15\% \leq x < 30\%$	15%	22.5%	19	427.5	64
6	$30\% \leq x < 45\%$	15%	37.5%	12	450	76
7	$45\% \leq x < 60\%$	15%	52.5%	4	210	80
8	$60\% \leq x < 75\%$	15%	67.5%	4	270	84
				$\Sigma f_i = 84$	1230	

$$\overline{X} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i} = \frac{1,230}{84} = 14.6$$

分組資料中央趨勢的衡量

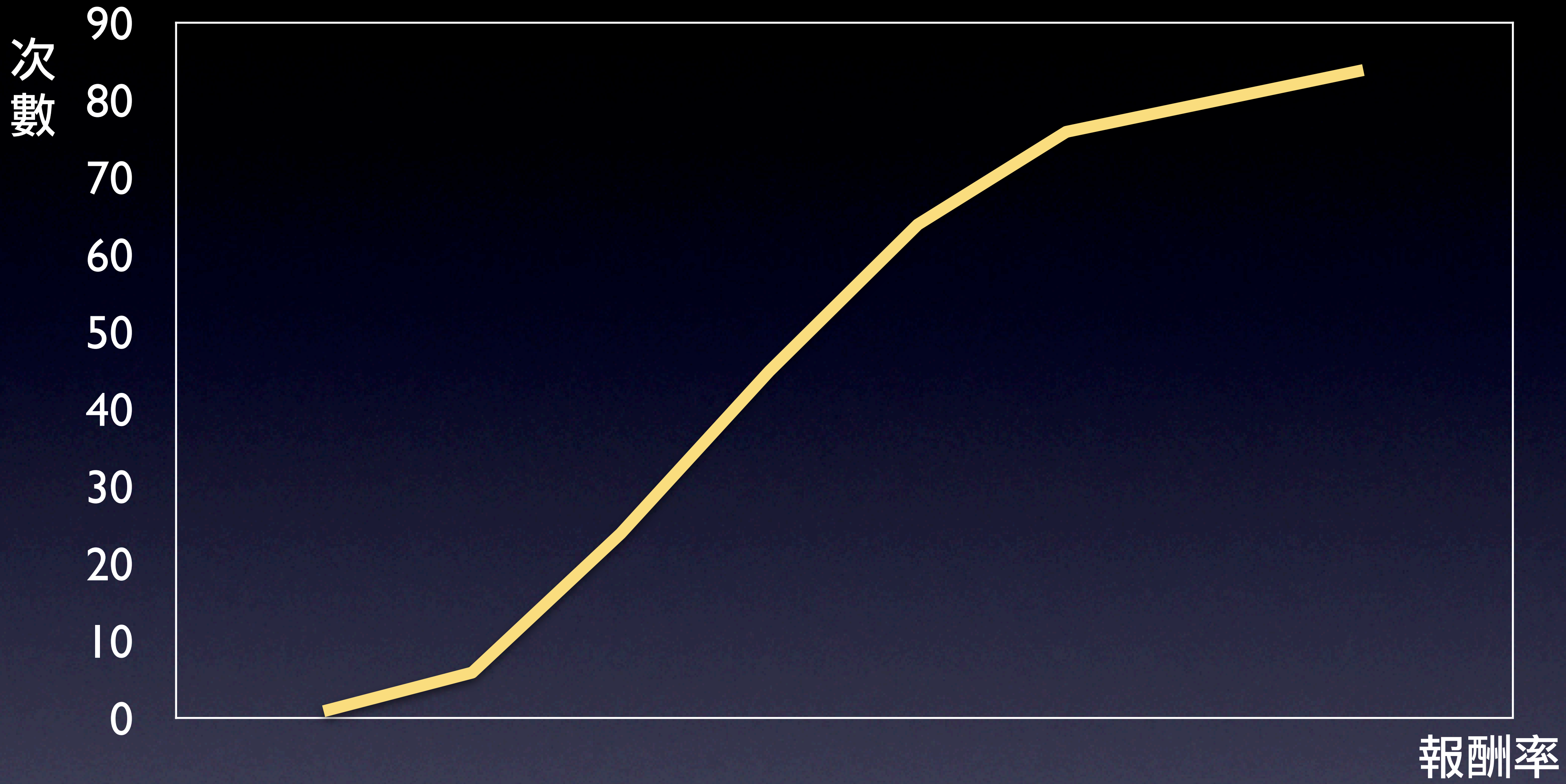
- 算術平均數

母體均數：$\mu = \dfrac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$　樣本均數：$\overline{X} = \dfrac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$

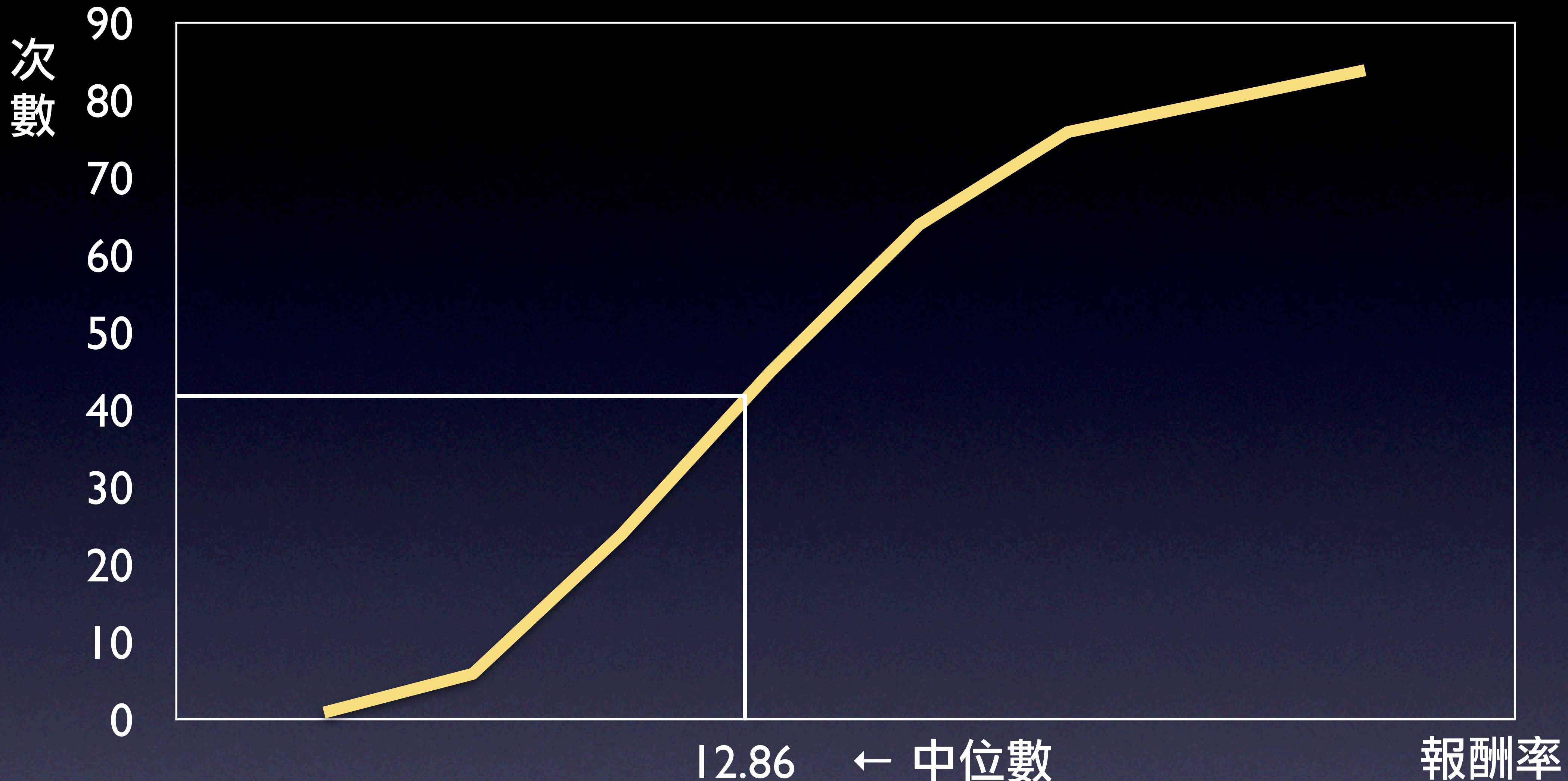
- 中位數

$$m_e = L_{m_e} + W_{m_e}\left(\frac{\frac{n}{2} - F_L}{f_{m_e}}\right)$$

股票基金近三年報酬率中位數的圖解



股票基金近三年報酬率中位數的圖解



$$m_e = L_{m_e} + W_{m_e}\left[\frac{\frac{n}{2} - F_L}{f_{m_e}}\right] = 0 + 15 \times \left[\frac{\frac{84}{2} - 24}{21}\right] = 0 + 12.86 = 12.86$$

分組資料中央趨勢的衡量

- **眾數**

  粗略法眾數

$$m_0 = \frac{(\text{組上界} + \text{組下界})}{2}$$

分組資料分散度的衡量

- 母體變異數與標準差

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{k}(x_i - \mu)^2 f_i \qquad \sigma = \sqrt{\sigma^2}$$

分組資料分散度的衡量

- 母體變異數與標準差

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{k}(x_i-\mu)^2 f_i \qquad \sigma = \sqrt{\sigma^2}$$

- 樣本變異數與標準差

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{k}(x_i-\overline{X})^2 f_i \qquad S = \sqrt{S^2}$$

股票型基金近三年報酬率的變異數與標準差

組別	組限	組中點	f_i 次數	$x - \overline{X}$	$(x - \overline{X})^2 f_i$
1	-45% ≤ x < -30%	-37.5%	1	-52.1	2714.41
2	-30% ≤ x < -15%	-22.5%	5	-37.1	6882.05
3	-15% ≤ x < 0%	-7.5%	18	-22.1	8791.38
4	0% ≤ x < 15%	7.5%	21	-7.1	1053.61
5	15% ≤ x < 30%	22.5%	19	7.9	1185.79
6	30% ≤ x < 45%	37.5%	12	22.9	6292.92
7	45% ≤ x < 60%	52.5%	4	37.9	5745.64
8	60% ≤ x < 75%	67.5%	4	52.9	11193.64
			$\Sigma = 84$		43864.44

股票型基金近三年報酬率的變異數與標準差

組別	組限	組中點	f_i 次數	$x-\overline{X}$	$(x-\overline{X})^2 f_i$
1	$-45\% \le x < -30\%$	-37.5%	1	-52.1	2714.41
2	$-30\% \le x < -15\%$	-22.5%	5	-37.1	6882.05
3	$-15\% \le x < 0\%$	-7.5%	18	-22.1	8791.38
4	$0\% \le x < 15\%$	7.5%	21	-7.1	1053.61
5	$15\% \le x < 30\%$	22.5%	19	7.9	1185.79
6	$30\% \le x < 45\%$	37.5%	12	22.9	6292.92
7	$45\% \le x < 60\%$	52.5%	4	37.9	5745.64
8	$60\% \le x < 75\%$	67.5%	4	52.9	11193.64
			$\Sigma = 84$		43864.44

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{k}(x_i - \overline{X})^2 f_i$$

$$= 43,864.44/(84-1) = 528.487$$

$$S = \sqrt{S^2} = \sqrt{528.487} = 22.99$$

未分組資料位置的衡量（其他測量數）

- 四分位數
  - 第 1 四分位數

$$Q_1 = L_{Q_1} + \frac{\frac{1}{4}n - F_{Q_1}}{f_{Q_1}} W_{Q_1}$$

未分組資料位置的衡量（其他測量數）

- 四分位數
  - 第 1 四分位數

$$Q_1 = L_{Q_1} + \frac{\frac{1}{4}n - F_{Q_1}}{f_{Q_1}} W_{Q_1}$$

  - 第 3 四分位數

$$Q_3 = L_{Q_3} + \frac{\frac{3}{4}n - F_{Q_3}}{f_{Q_3}} W_{Q_3}$$

學生英文考試成績次數分配表

組號	組限	次數	累加次數
1	30~40	2	2
2	40~50	1	3
3	50~60	12	15
4	60~70	14	29
5	70~80	38	67
6	80~90	33	100
7	90~100	6	106

學生英文考試成績次數分配表

組號	組限	次數	累加次數
1	30~40	2	2
2	40~50	1	3
3	50~60	12	15
4	60~70	14	29
5	70~80	38	67
6	80~90	33	100
7	90~100	6	106

$$Q_1 = L_{Q_1} + \frac{\frac{1}{4}n - F_{Q_1}}{f_{Q_1}} W_{Q_1} = 60 + \frac{26.5 - 15}{14} \times 10 = 68.21$$

學生英文考試成績次數分配表

組號	組限	次數	累加次數
1	30~40	2	2
2	40~50	1	3
3	50~60	12	15
4	60~70	14	29
5	70~80	38	67
6	80~90	33	100
7	90~100	6	106

$$Q_1 = L_{Q_1} + \frac{\frac{1}{4}n - F_{Q_1}}{f_{Q_1}} W_{Q_1} = 60 + \frac{26.5 - 15}{14} \times 10 = 68.21$$

$$Q_3 = L_{Q_3} + \frac{\frac{3}{4}n - F_{Q_3}}{f_{Q_3}} W_{Q_3} = 80 + \frac{\frac{3}{4} \times 106 - 67}{33} \times 10 = 83.79$$

未分組資料位置的衡量（其他測量數）

- 十分位數

$$D_i = L_{D_i} + \frac{\frac{n \cdot i}{10} - F_{D_i}}{f_{D_i}} W_{D_i}$$

未分組資料位置的衡量（其他測量數）

- 十分位數

$$D_i = L_{D_i} + \frac{\frac{n \cdot i}{10} - F_{D_i}}{f_{D_i}} W_{D_i}$$

- 百分位數

$$P_i = L_{P_i} + \frac{\frac{n \cdot i}{100} - F_{P_i}}{f_{P_i}} W_{P_i}$$

學生英文考試成績次數分配表

組號	組限	次數	累加次數
1	30~40	2	2
2	40~50	1	3
3	50~60	12	15
4	60~70	14	29
5	70~80	38	67
6	80~90	33	100
7	90~100	6	106

學生英文考試成績次數分配表

組號	組限	次數	累加次數
1	30~40	2	2
2	40~50	1	3
3	50~60	12	15
4	60~70	14	29
5	70~80	38	67
6	80~90	33	100
7	90~100	6	106

$$D_6 = L_{D_6} + \frac{\frac{n \cdot 6}{10} - F_{D_6}}{f_{D_6}} W_{D_6} = 70 + \frac{\frac{106 \times 6}{10} - 29}{38} \times 10 = 79.1$$

學生英文考試成績次數分配表

組號	組限	次數	累加次數
1	30~40	2	2
2	40~50	1	3
3	50~60	12	15
4	60~70	14	29
5	70~80	38	67
6	80~90	33	100
7	90~100	6	106

$$D_6 = L_{D_6} + \frac{\frac{n \cdot 6}{10} - F_{D_6}}{f_{D_6}} W_{D_6} = 70 + \frac{\frac{106 \times 6}{10} - 29}{38} \times 10 = 79.1$$

$$P_i = L_{P_i} + \frac{\frac{n \cdot i}{100} - F_{P_i}}{f_{P_i}} W_{P_i} = 90 + \frac{\frac{106 \times 95}{10} - 100}{6} \cdot 10 = 90 + 1.167 = 91.167$$

Excel 的使用

- 利用 Excel 求算各種統計測量數

1. 開啟 Excel 工作表，輸入 84 種股票型基金的報酬率。

Excel 的使用

- 利用 Excel 求算各種統計測量數
  1. 開啟 Excel 工作表，輸入 84 種股票型基金的報酬率。
  2. 選取「工具」→「資料分析」→「敍述統計」，按確定。

Excel 的使用

- 利用 Excel 求算各種統計測量數

1. 開啟 Excel 工作表，輸入 84 種股票型基金的報酬率。
2. 選取「工具」→「資料分析」→「敍述統計」，按確定。

	A	B	C
1	股票型基金的報酬率		
2			
3	平均數	14.55511905	
4	標準誤	2.445867389	
5	中間值	13.87	中位數
6	眾數	#N/A	
7	標準差	22.4167449	
8	變異數	502.5104518	
9	峰度	-0.004619895	
10	偏態	0.364213066	
11	範圍	110.33	全距
12	個數	84	樣本數