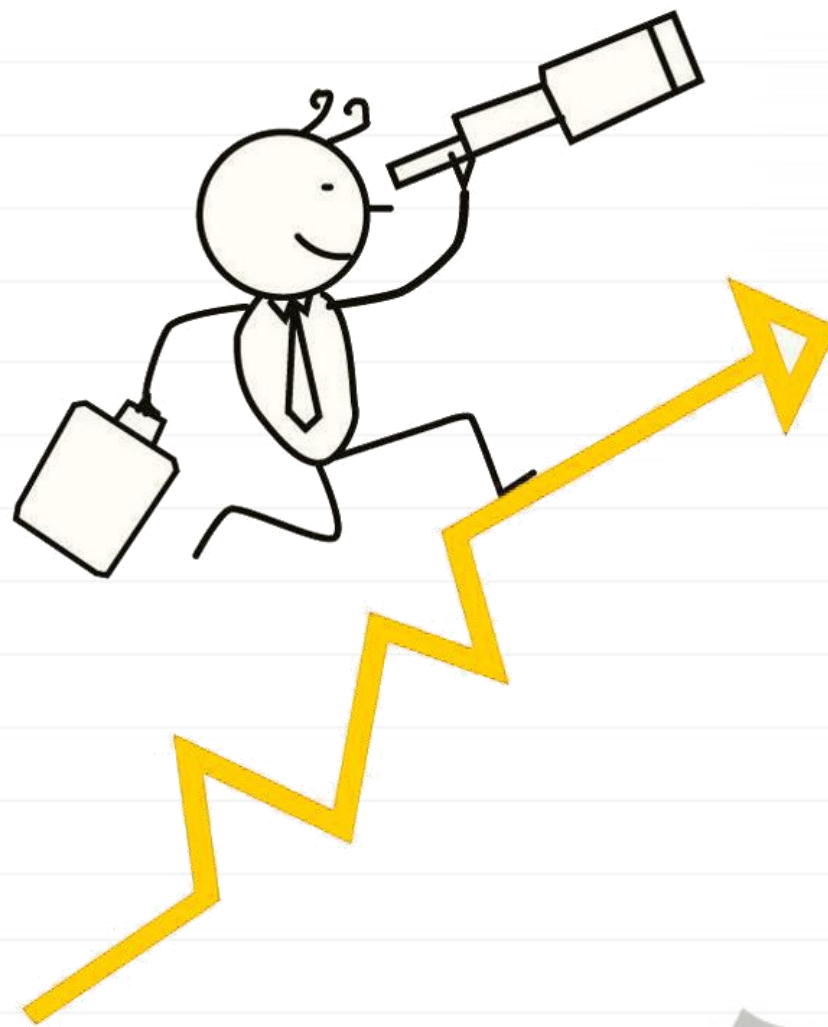


統計就是要這樣跑 CH5-7

分享者：張芳全教授



目錄

1



集中量數

3



相對地位量數

2



變異量數

4

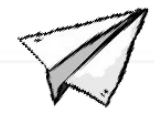


Q & A



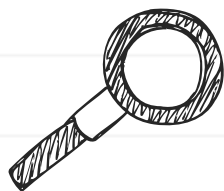
集中量數

- 01 幾何平均數
- 02 中位數
- 03 眾數



集中量數

可計算某一變項的數值集中情形。



算術平均數(平均數)

average mean, M ，若由 X 變數計算，記為 X bar, \bar{X}



幾何平均數

一種由 n 個正數之乘積的 n 次根表示的平均數



中數

median, M_d / M_{dn}



眾數

mode, M_o



請問下列數列中，

1、平均數為？

2、中數為？

3、眾數為？

2、3、3、3、5、8



平均數 M

- (一) 未分組資料計算平均數
- (二) 加權平均數

p. 100-101

特性：

- 1、一組資料中，每個變數與平均數之差(離均差)為0
- 2、每一個數+常數C，新平均數=原平均數+常數C
- 3、每一個數 \times 常數C，新平均數=原平均數 \times 常數C



- (1) 反應靈敏，反映資料中任何一個數值的大小變化。
- (2) 計算嚴密，確定公式，同一組數據任何人算都一樣。
- (3) 計算簡單。
- (4) 概念容易理解。
- (5) 較少受抽樣影響，樣本大小或個體變化影響不大。



- (1) 易受極端值影響。
- (2) 模糊不清的數據，無法計算平均數。



中數 Md/Mdn

- (一) 數據個數奇數
- (二) 數據個數偶數

p. 102-103

特性：

能描述一組資料的典型情況。在這組資料中有一半的數據比它大，有一半的數據比它小。



- (1) 計算簡單。
- (2) 極端值變化不影響。

例如：

3、5、26、50、1000
3、5、26、600、2000



- (1) 受抽樣影響。
- (2) 反應不夠靈敏。3、5、26、50、300
- (3) 不能做代數運算。
- (4) 不太普遍運用。



眾數 Mo

p. 103-104

特性：

能描述一組資料的集中趨勢。在這組資料中出現最多次數的數值。



- (1) 概念簡單。
- (2) 容易理解。
- (3) 較少受極端值影響。



- (1) 反應不靈敏。
- (2) 不能做代數運算。
- (3) 不太普遍運用。



次數分配、百分比

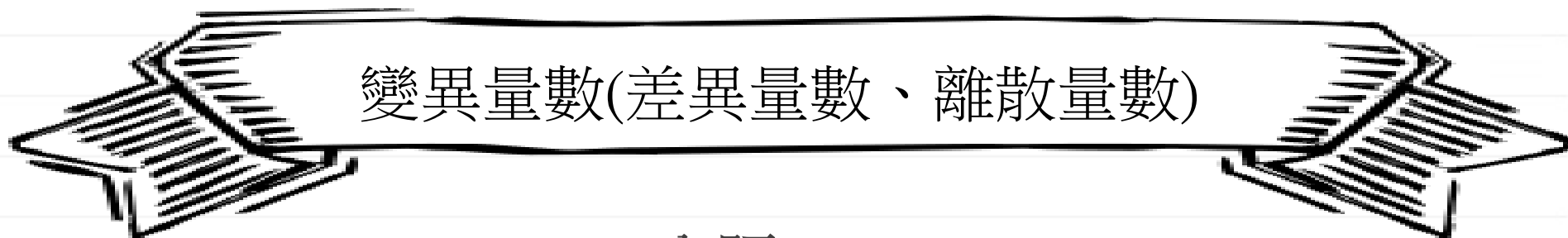
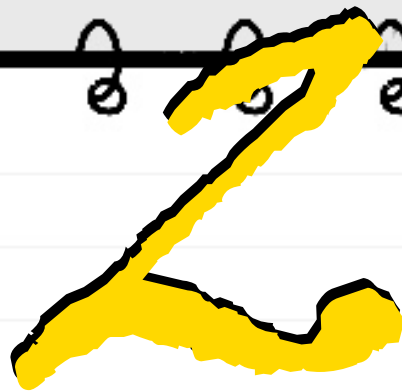
分析-敘述統計-次數分配表-放變數(性別、學歷)-確定-看報表

平均數、標準差

分析-敘述統計-敘述統計-放變數(影1-影8)-選項-勾平均值和標準差-繼續-確定-看報表

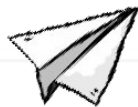
注意

整本論文的小數點位數統一較佳。



變異量數(差異量數、離散量數)

- 01 全距
- 02 平均差
- 03 標準差



變異量數

可瞭解某一變項的數值分散程度。

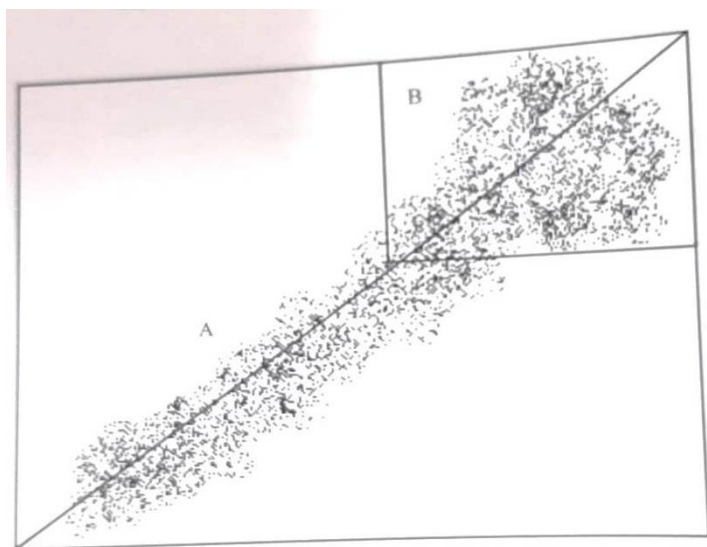


圖 6-1 變異量數示意圖

A 的樣本群變異小
B 的樣本群變異大



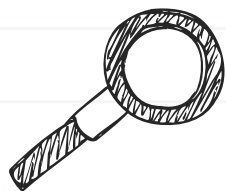
全距
range, R



平均差
average deviation, AD/MD



標準差(絕對差量)
standard deviation, S/SD



四分位差



全距 R 說明資料離散程度的統計量

p. 119

特性：

最簡單最易懂，但也最粗糙不可靠，易受極端值影響。

2、3、3、3、5、8

$$R = 8 - 2 = 6$$



平均差 MD/AD p. 121-122

特性：

1、能較好反應次數分配的離散程度。(與平均數的距離，有正負→離均差，取絕對值加總的平均值→平均差。

2、所有資料與平均數的絕對離差的平均值。

=所有資料與平均數的差之絕對值/全部資料的個數

$$= \frac{\sum |X_i - \bar{X}|}{n}$$

3、能反應數據的集中程度，較少用於推論統計分析。



標準差 S/SD 與 變異數 S^2

p. 122-127

特性：
$$\frac{\sum (X_i - \bar{X})^2}{N} = S^2$$

1、符號

(1) 母群參數標準差 σ 與 變異數 σ^2

(2) 樣本統計量標準差 s 與 變異數 s^2 (分母 $N-1$ ，避免低估)

2、表示全部資料分配的離散程度。標準差大 v.s 標準差小

3、每一個數 + 常數 C ，新標準差 = 原標準差

4、每一個數 \times 常數 C ，新標準差 = 原標準差 \times 常數 C

5、每一個數 \times 常數 C + 常數 D ，新標準差 = 原標準差 \times 常數 C



差異係數 CV

p. 126-127

特性：

- 1、用在觀測工具不同，例如身高和體重，以及樣本間水準差異大，例如40人和4000人的數據。
- 2、資料須具有等距尺度。
- 3、變項具有絕對零點。
- 4、不能進行統計推論。

$$CV = \frac{S}{\bar{X}} \times 100 \%$$

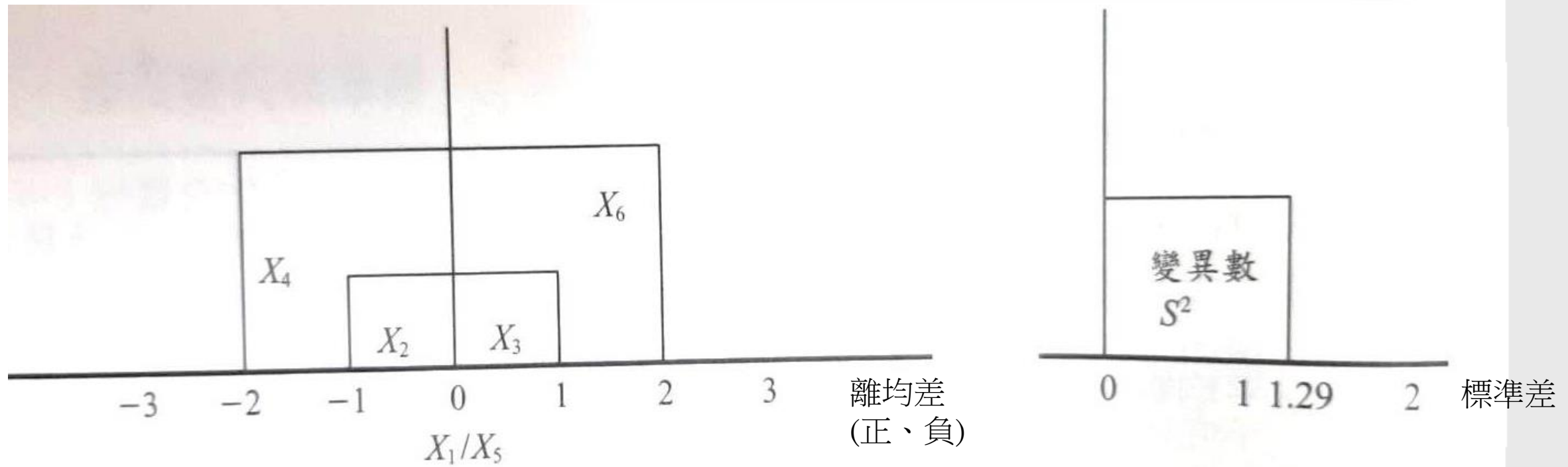
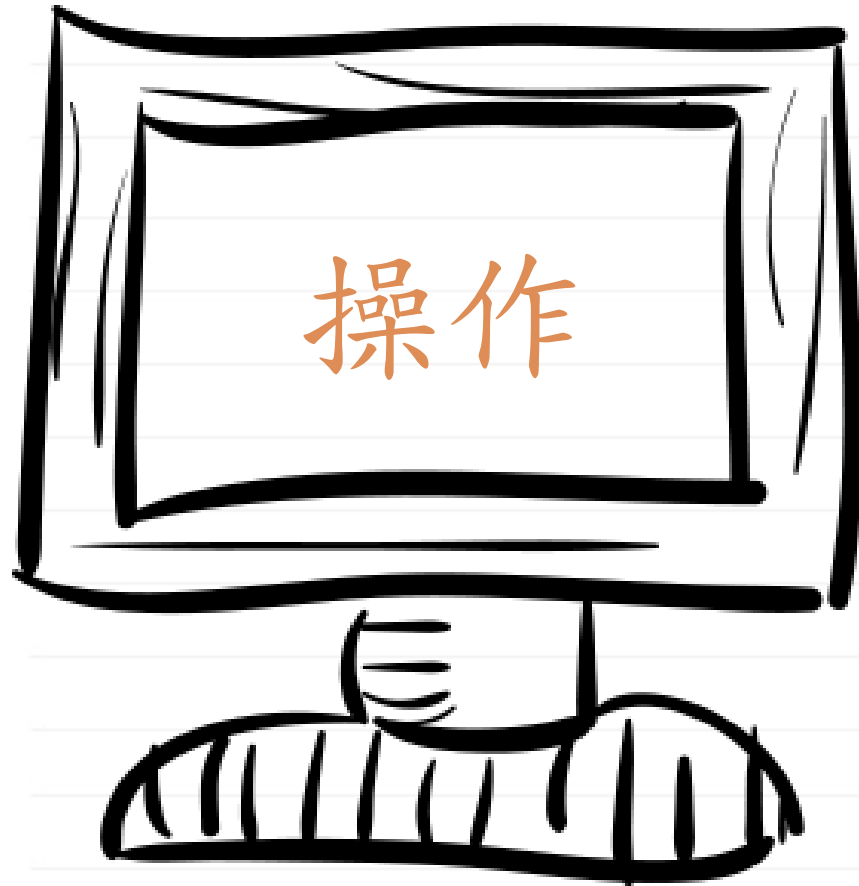


圖 6-3 離均差、變異數、標準差之關係



平均數、標準差

分析-敘述統計-次數分配表-放變數(過3之1-10)-勾選標準差、變異、範圍-確定-看報表

差異係數計算

spss資料複製-excel貼上-鍵入方程式-確定-下拉-看報表



動差體系

p. 119-120

以平均數為原點的動差，稱為中心動差。常見的中心動差有四級。

一級動差 $\mu_1 = \frac{\Sigma(X_i - \bar{X})}{N} = 0$ 1次方，分子常會取絕對值總和→平均差。

二級動差 $\mu_2 = \frac{\Sigma(X_i - \bar{X})^2}{N} = S^2$ 2次方，變異數。

三級動差 $\mu_3 = \frac{\Sigma(X_i - \bar{X})^3}{N}$ 3次方，偏態或偏斜度的指標。(正+負-常0)

四級動差 $\mu_4 = \frac{\Sigma(X_i - \bar{X})^4}{N}$ 4次方，峰態性指標。(大-高狹；小-低闊)
 $K = \mu_4 / \mu_2^2 - 3$ ， $K > 0$ 高狹 $K < 0$ 低闊 $K = 0$ 常態

Kline(1998)偏態絕對值 < 3 ，峰度絕對值 < 10 →常態分配。

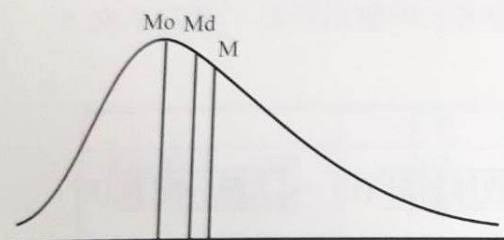
常態分配之 Skewness 為 0，Kurtosis 為 0。



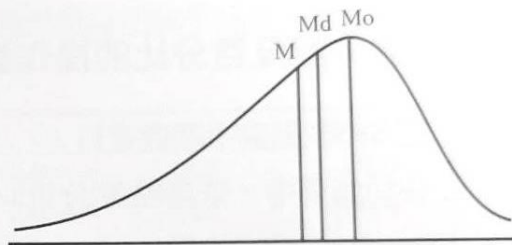
偏態和峰度

p. 104-105、p. 129-130

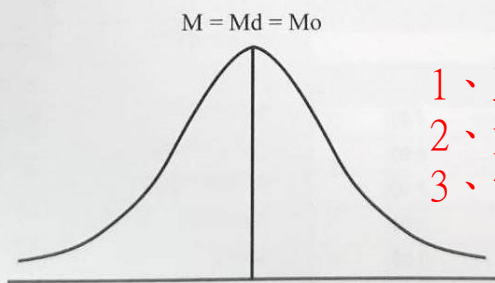
表示一組量數的分配情形，偏態分為正偏態、負偏態、常態分配。
三種態勢上平均數、中數和眾數落點不同。



正偏態



負偏態



常態分配

- 1、正偏態： $M > Md > Mo$
- 2、負偏態： $M < Md < Mo$
- 3、常態分配： $M = Md = Mo$

圖 5-1 不同分配中三個集中量數的關係

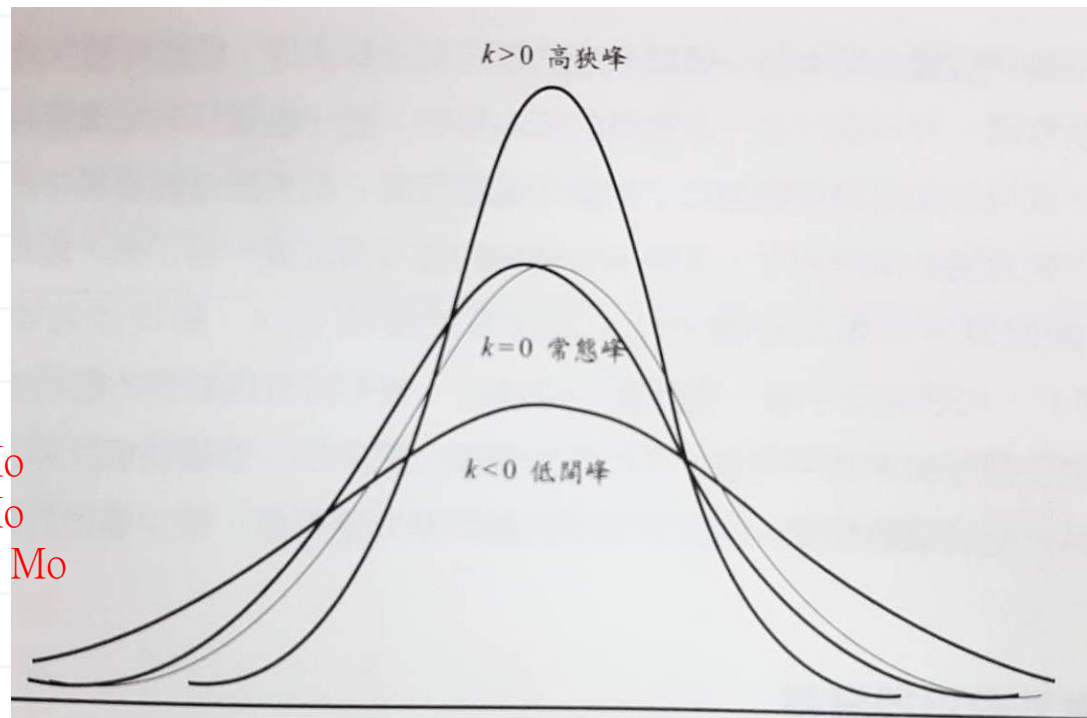
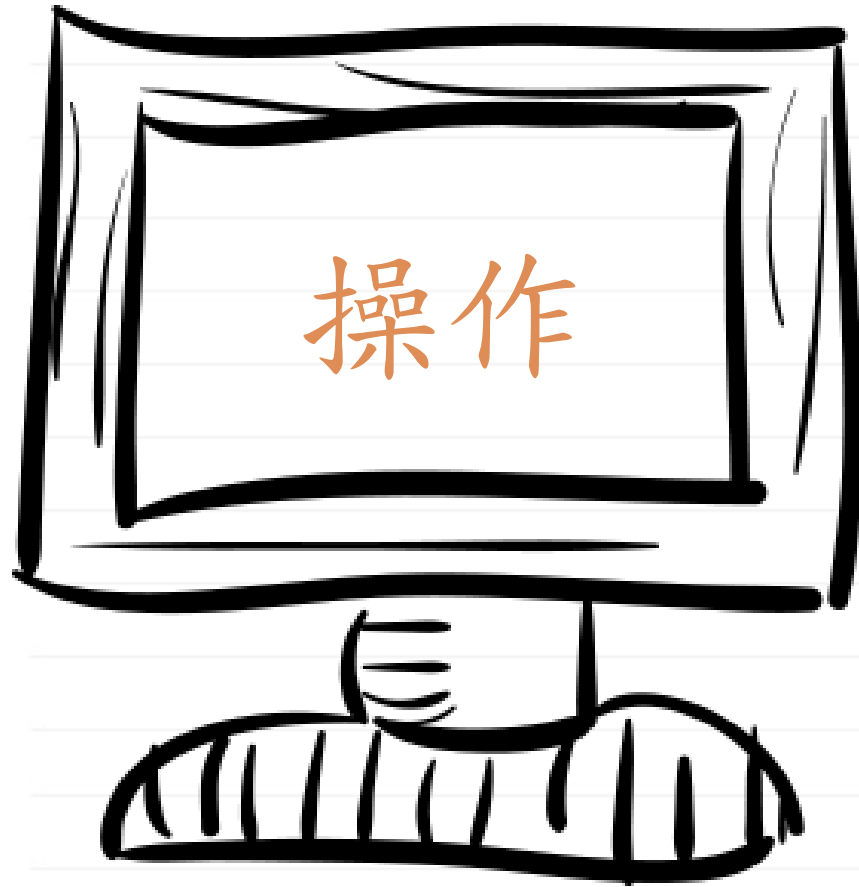


圖 6-2 高狹峰、常態峰與低闊峰之示意圖



分配圖

分析-敘述統計-次數分配表-放變數(年資、職務)-**圖表-勾選直方圖+在直方圖上顯示常態曲線**-繼續-確定-看報表

峰度、偏態

分析-敘述統計-次數分配表-放變數(年資、職務)-**統計資料-勾選偏態和峰值**-確定-看報表

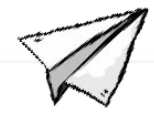


相對地位量數

可瞭解某一個體在團體中的相對地位。

01 百分等級(percentile rank, PR)

02 標準化Z分數(標準分數/Z score)





百分等級 PR

p. 131-134

- 1、符號： $P_1 = 5.92$ ，第一個百分位數的人得分為5.92
- 2、最高百分等級為PR99，最低為PR1，表示在100人中勝過99人或是1人。
- 3、屬於次序變數，無法用代數計算。

$$P_p = L + \left[\frac{\frac{PR}{100} \times N - F}{f_p} \right] \times i$$

式中，PR 為百分等級；

L 為第 p 百分位數所在組的精確下限；

f_p 為第 p 百分位數所在組的次數；

F 為第 p 百分位數所在組以下的累積次數；

N 為總次數；

i 為組距。

精確下限的意義是指數字（該組最低數值者）減 0.5，這是約定成俗。

$$PR = \frac{100}{N} \times \left[F + \frac{f_b (X - L)}{i} \right]$$

式中， X 為原始分數；

f_b 為該原始分數所在組的次數；

L 為該原始分數所在組的精確下限；

F 為小於 L 的累積次數；

N 為總次數；

i 為組距。



標準分數(Z分數) p. 134-138

特性：

1、以標準差為單位表示一個原始分數在團體中的相對位置量數。

2、公式 $Z = \frac{X - \bar{X}}{S} = \frac{x}{s}$

3、來算吧：某班平均成績為90分，標準差3分，甲生得96分，乙生得93分，請問甲乙生的Z分數各為多少？

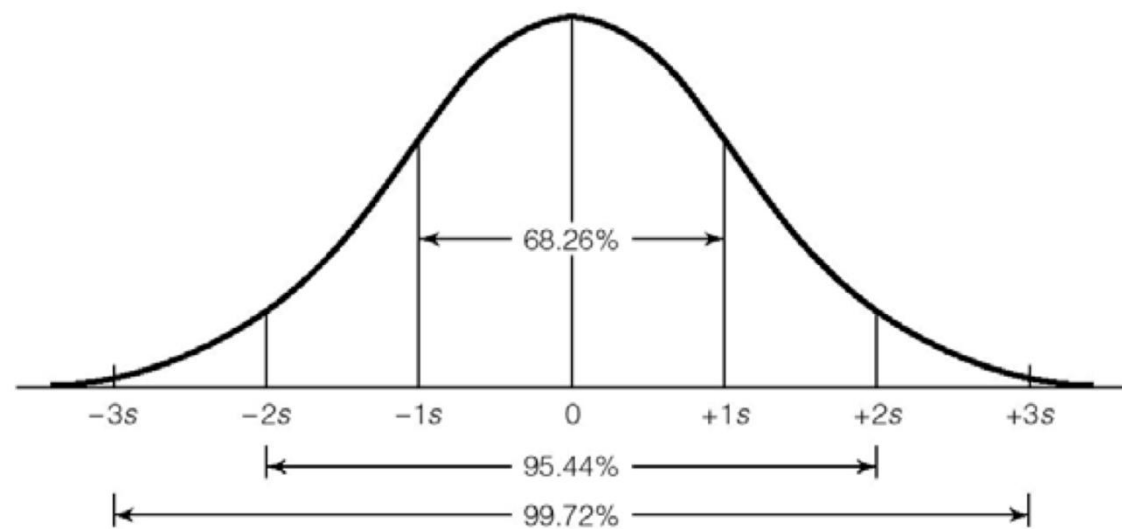
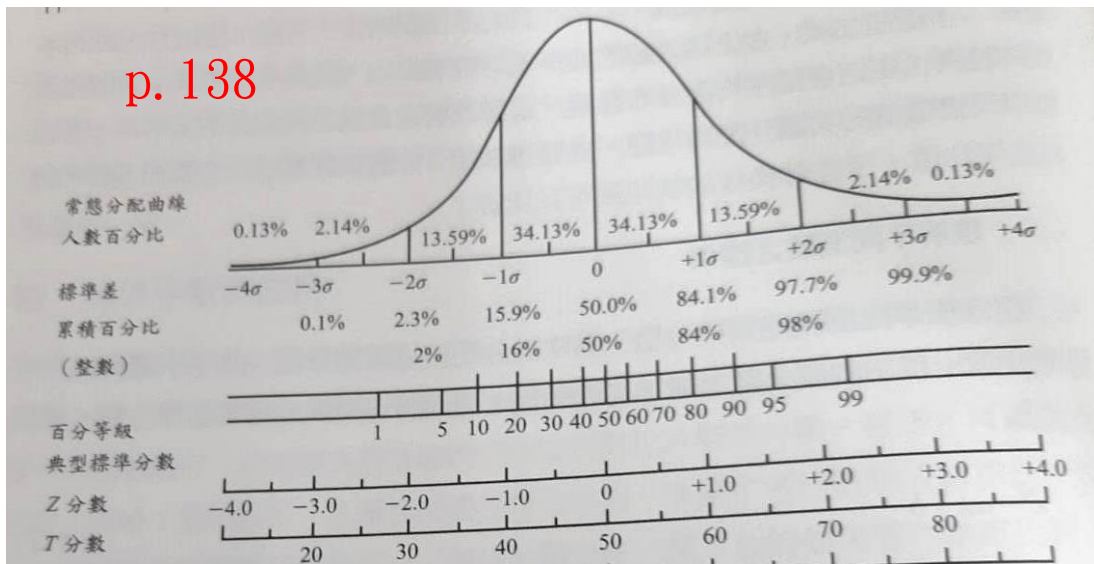
4、Z分數無實際單位。

5、標準常態分配：平均值0，標準差1。

6、可以比較、可以相加、明確性(常態分配查表)、穩定性。

7、Z分數有0和負值，常用的是T分數： $T=10Z+50$ 。

p. 138





常態分配 p. 138-140

平均數在0以下，佔全體人數的一半(50%)。

平均數在0以上，佔全體人數的一半(50%)。

得分在 $\mu \pm 1$ 個標準差(σ)內，佔總人數68.26%。

得分在 $\mu \pm 2$ 個標準差(σ)內，佔總人數95.44%。

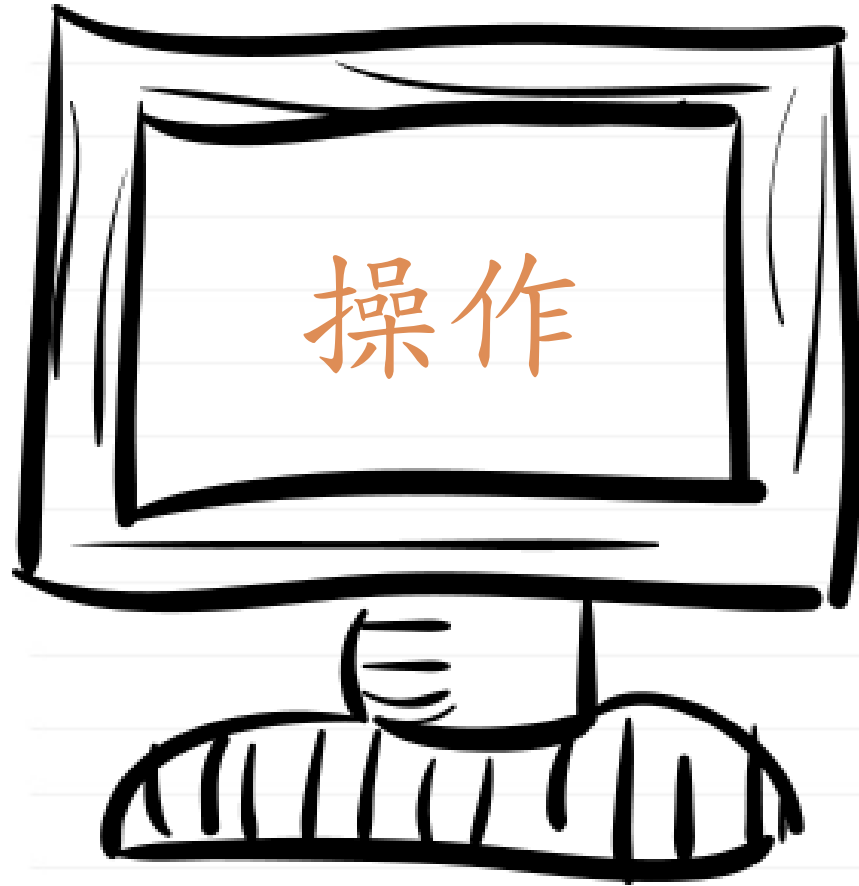
得分在 $\mu \pm 3$ 個標準差(σ)內，佔總人數99.72%。

若未剛好落在1、2、3個標準差上，需查附表D。

練習查表：

$$Z=2.3, p = ?$$

$$Z= -1.2, p = ?$$



百分位數

轉換-秩觀察值-放變數(爭議加總)-秩類型-勾選等級、分數秩以%表示-繼續-確定-回主畫面看報表

標準分數

分析-敘述統計-敘述統計-放變數(爭議加總)-勾選將標準化值存成變數-確定-看報表

感謝聆聽

