統 計 學

第三章 統計資料之分析解釋(一)

3-1 統計分析與統計量數
3-2 平均數
3-3 算術平均數
3-4 幾何平均數
3-5 調和平均數(選讀)
3-6 中位數及其他分位數
3-7 眾數
3-8 分組資料平均數之計算
3-9 各種平均數之關係與比較

書號：512282

編著 江建良

普林斯頓國際有限公司

3-1 統計分析與統計量數

- **統計分析的意義：**

  統計分析，係指求算一些統計數值來表達統計資料的特徵，以了解資料的特性。

- **統計分析的方法：**
  - **時間數列：** 常見之分析方法為時間數列分析與指數。
  - **空間數列：** 比照屬性數列之分析方法。
  - **屬性數列：** 求算比例、百分比或獨立性檢定分析。
  - **屬量數列：** 常見之分析方法有次數分配、變異數分析、相關與迴歸等。

- **統計量數的種類：**
  - **平均數**
  - **差異量數**
  - **偏態係數**
  - **峰態係數**

3-2 平均數

平均數的意義：

1. 是次數分配的中心位置，因此又稱為地位量數(measure of location)。
2. 大多數平均數的計算，都是由“平均方式”而得，因此稱為平均數。
3. 就大自然現象觀察，發現各種現象均有向其中心集中的趨勢，因此平均數又稱為集中趨勢量數(central tendecy measures)。

平均數的功用：

- 簡化作用
- 代表作用
- 比較作用

平均數的種類：

- 平均數
- 中位數
- 眾數

3-3 算術平均數

- 算術平均數的計算方式：

  - 母體資料：

$$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i = \frac{1}{N}\sum X = \frac{1}{N}(x_1 + x_2 + \cdots + x_N)$$

  - 樣本資料：

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\sum X = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

例題 3-1

茲就下列10個數值資料（設爲母體資料），求其算數平均數

15 , 10 , 12 , 8 , 6 , 25 , 18 , 5 , 20 , 11

解：

$$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i = \frac{1}{N}(x_1 + x_2 + \cdots\cdots + x_N)$$

$$= \frac{1}{10}(15 + 10 + \cdots\cdots + 20 + 11) = 13$$

例題 3-2

設有兩組學生統計學成績如下（設爲樣本資料），試問何組爲優？

A組：90, 80, 75, 40, 45分

B組：85, 70, 60, 50, 45分

解：A組：$\overline{X}_A = \frac{1}{n}\sum X = \frac{1}{5}(90 + 80 + 75 + 40 + 45) = 66$（分）

B組：$\overline{X}_B = \frac{1}{n}\sum X = \frac{1}{5}(85 + 70 + 60 + 50 + 45) = 62$（分）

由上計算可知，A組較B組爲優(66分大於62分)。

例題 3-3

甲生各科成績及每週上課時數如下：

科　目	成績 (分)	每週上課時數
國　文	80	5
英　文	70	4
會計學	75	5
統計學	85	3

試求算 (a) 加權平均數。

(b) 簡單平均數。

解： (a) 加權平均數 $=\dfrac{\sum \omega X}{\sum \omega}=\dfrac{5\times 80+4\times 70+5\times 75+3\times 85}{5+4+5+3}=\dfrac{1{,}310}{17}$

$=77.06$ (分)

(b) 簡單平均數 $=\dfrac{1}{n}\sum X=\dfrac{80+70+75+85}{4}=\dfrac{310}{4}=77.5$ (分)

例題 3-4

甲班學生50人分爲四組，其統計學成績平均分數如下：

A組：15人，平均分數70分

B組：15人，平均分數80分

C組：10人，平均分數65分

D組：10人，平均分數60分

試由上述分組資料，求算甲班50人統計學平均分數。

解：

$$\mu=\frac{15\times 70+15\times 80+10\times 65+10\times 60}{15+15+10+10}=\frac{3{,}500}{50}=70\,(\text{分})$$

例題 3-5

試求2, 2, 4, 4, 100之算術平均數

解： $\mu = \dfrac{2+2+4+4+100}{5} = 22.4$

由於100屬極端值，若計入則平均數爲22.4，與資料顯然不合，故應將100剔除，再計算2, 2, 4, 4等四個之平均數 $\left(\mu = \dfrac{2+2+4+4}{4} = 3\right)$ 較爲合理。

3-4 幾何平均數

- 幾何平均數的計算方式：

  - 母體資料：

$$G = \sqrt[N]{x_1 \times x_2 \times \cdots \times x_N} = \sqrt[N]{\pi x}$$

  - 樣本資料：

$$G = \sqrt[n]{x_1 \times x_2 \times \cdots \times x_n} = \sqrt[n]{\pi x}$$

例題 3-6

台泥公司最近四年之營業額為3, 6, 12, 24 (億元)，試求其每年平均營業額。

解：因數列成幾何級數 (即等比級數)，所以用幾何平均數求之。

$$G = \sqrt[4]{3 \times 6 \times 12 \times 24} = 8.49 \text{ (億元)}$$

例題 3-7

李經理第1年投資聯電股票每股市價50元，第1年底聯電股票每股市價60元，而第2年底聯電股票每股市價50元，試求李經理兩年度投資聯電股票的總報率與每年平均報酬率。

解：(a) 兩年度總報率：

第1年報酬率 $= (\$60 - 50) \div 50 = 20\%$

第2年報酬率 $= (\$50 - 60) \div 60 = -16.67\%$

兩年度總報率 $= 20\% + (-16.67\%) = 3.33\%$

(b) 平均每年報酬率：

① 算術平均數 $=[120\%+(100\%-16.67\%)]\div 2=101.67\%$

② 幾何平均數 $=G=\sqrt[n]{(1+X_1)\times(1+X_2)\times\cdots\times(1+X_n)}$

$$=\sqrt[2]{(1+20\%)\times(1-16.67\%)}=100\%$$

由於第1年漲10元，第2年跌10元，因此李經理持有兩年的報酬率為100%（表示無變動，若以算術平均數計算，上漲1.67%），以幾何平均數計算的結果比算術平均數計算的結果較為精確。

註：由於報酬率計算是以100% 為比較基準（避免為負數），故上漲20%，須以120% 表示 (1+20%)，下跌16.67%，須以83.33% 表示。

3-5 調和平均數（選讀）

- 調和平均數之計算方式：

$$H = \frac{n}{\frac{1}{X_1} + \frac{1}{X_2} + \cdots \frac{1}{X_n}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}} = \frac{n}{\sum \frac{1}{X}}$$

例題 3-8

(a) 甲生返往兩地，去時每小時平均4公里，回程每小時平均5公里，試求往返兩程之平均速率？

(b) 乙生每次購買雞蛋四個，第1日蛋價每元4個，第2日每元5個，第3日每元2個，試求每元平均買幾個蛋？

解：(a) 往返兩地之距離相等，且速率以距離 / 時間表示 (分子固定)。

$$H=\frac{X+X}{\frac{X}{4}+\frac{X}{5}}=\frac{X}{X}\times\frac{40}{9}=\frac{40}{9}\text{ (公里 / 小時)(設每段距離 }X\text{)}$$

說明：設去程與回程均 x 公里，總共 $2x$ 公里，而共花 $\left(\frac{x}{4}+\frac{x}{5}\right)$ 小時，若平均速率以算術平均數 $\frac{4+5}{2}=4.5$ 公里 / 小時計算，$\left(\frac{x}{4}+\frac{x}{5}\right)$ 小時共走 $\left(\frac{x}{4}+\frac{x}{5}\right)\times 4.5=2.025x$ 公里，顯然不對，但若以調和平均數計算，則 $\left(\frac{x}{4}+\frac{x}{5}\right)$ 小時共走 $\left(\frac{x}{4}+\frac{x}{5}\right)\times\frac{40}{9}=2x$ 公里，以題目吻合，故平均速率應以調和平均數計算。

(b) 每次購買四個，且物價以數量 / 價值表示 (分子固定)。

$$H=\frac{4+4+4}{\frac{4}{4}+\frac{4}{5}+\frac{4}{2}}=\frac{4}{4}\times\frac{1+1+1}{\frac{1}{4}+\frac{1}{5}+\frac{1}{2}}=3.16\text{ (個 / 元)}$$

中位數的計算方式：

1. n 為奇數時，

   中位數在第 $\frac{n}{2}+\frac{1}{2}$ 項。

2. n 為偶數時，

   中位數在第 $\frac{n}{2}$ 項與 $\frac{n}{2}+1$ 項之算術平均數。

例題 3-9

求數列8, 6, 4, 5, 3, 10, 12之中位數。

解：(a) 先排列順序 (從小到大)

$$3,4,5,6,8,10,12$$

(b) 由於是奇數個 (n = 7)，所以中位數位於第 $\frac{7}{2}+\frac{1}{2}=4$ 項之位置， 所以中位數為 6。

例題 3-10

求數列8, 6, 4, 5, 3, 10之中位數。

解：(a) 先排列順序 (從小到大)

$$3,4,5,6,8,10$$

(b) 由於是偶數個，中位數介於第 $\frac{n}{2}=\frac{6}{2}=3$ 項與第 $\frac{n}{2}+1=\frac{6}{2}+1=4$ 項之間。

(c) 第3項為5，第4項為6，故中位數 $M_e=\frac{5+6}{2}=5.5$。

3-6 中位數及其他分位數

其他分位數：

四分位數：

1. 先將個變量從小到大順序排列。

2. 從最小的一端起算。

(1) $Q_1 \Rightarrow$ 第 $\frac{n}{4}+\frac{1}{2}$ 項之數值。

(2) $Q_2 \Rightarrow M_e \Rightarrow$ 第 $\frac{n}{2}+\frac{1}{2}$ 項之數值。

(3) $Q_3 \Rightarrow$ 第 $\frac{3n}{4}+\frac{1}{2}$ 項之數值。

例題 3-11

求數列60, 62, 78, 76, 75, 84, 95, 90, 80之 Q_1 , Q_2 , Q_3 。

解：(a) 先按大小順序排列

60, 62, 75, 76, 78, 80, 84, 90, 95

(b) ① Q_1 在第 $\frac{9}{4}+\frac{1}{2}$ 項 $= 2.75$ 項

$\therefore \quad Q_1 = 62 + 0.75(75-62) = 71.75$

[註：0.75項即佔62到75之距離的75%]

② $Q_2 = M_e$ 在第 $\frac{9}{2}+\frac{1}{2} = 5$ 項

$\therefore \quad Q_2 = M_e = 78$

③ Q_3 在第 $\frac{3\times 9}{4}+\frac{1}{2} = 7.25$ 項

$\therefore \quad Q_3 = 84 + 0.25(90-84) = 85.5$

3-6 中位數及其他分位數

其他中位數：

十分位數：

(1) $D_1 \Rightarrow$ 第 $\dfrac{n}{10}+\dfrac{1}{2}$ 項之數值。

(2) $D_2 \Rightarrow$ 第 $\dfrac{2n}{10}+\dfrac{1}{2}$ 項之數值。

…………………………

(3) $D_9 \Rightarrow$ 第 $\dfrac{9n}{10}+\dfrac{1}{2}$ 項之數值。

即第 r 個十分位數 D_r 在第 $\dfrac{rn}{10}+\dfrac{1}{2}$ 項之數值。

3-6 中位數及其他分位數

其他中位數：

百分位數：

(1) $P_1 \Rightarrow$ 第 $\frac{n}{100}+\frac{1}{2}$ 項之數值。

(2) $P_2 \Rightarrow$ 第 $\frac{2n}{100}+\frac{1}{2}$ 項之數值。

…………………………

(3) $P_{99} \Rightarrow$ 第 $\frac{99n}{100}+\frac{1}{2}$ 項之數值。

即第 r 個百分位數在第 $\frac{rn}{100}+\frac{1}{2}$ 項之數值。

例題 3-12

試就上例求算：

(a) 第三個「十分位數」($\boldsymbol{D_3}$)及第八個「十分位數」($\boldsymbol{D_8}$)。

(b) 第四十個「百分位數」($\boldsymbol{P_{40}}$)及第六十個「百分位數」($\boldsymbol{P_{60}}$)。

解：(a) 先按大小順序排列

$$60,\ 62,\ 75,\ 76,\ 78,\ 80,\ 84,\ 90,\ 95$$

(b) ① (i) $\boldsymbol{D_3}$ 在 $\dfrac{3\times 9}{10}+\dfrac{1}{2}=3.2$ 項 $\therefore\quad D_3=75+0.2(76-75)=75.2$

(ii) $\boldsymbol{D_8}$ 在 $\dfrac{8\times 9}{10}+\dfrac{1}{2}=7.7$ 項

$$D_8=84+0.7(90-84)=88.2$$

② (i) $\boldsymbol{P_{40}}$ 在第 $\dfrac{40\times 9}{100}+\dfrac{1}{2}$ 項 $=4.1$ 項

$$P_{40}=76+0.1(78-76)=76.2$$

(ii) $\boldsymbol{P_{60}}$ 在第 $\dfrac{60\times 9}{100}+\dfrac{1}{2}$ 項 $=5.9$ 項

$$P_{60}=78+0.9(80-78)=79.8$$

例題 3-13

試求數列 5,8,3,9,5,6,8 之 Q_2 , Me , D_5 , P_{50} 。

解：(a) 先按大小順序排列

3 , 5 , 5 , 6 , 8 , 8 , 9

(b) $\boldsymbol{Q_2}$ 在第 $\dfrac{2\times 7}{4}+\dfrac{1}{2}=4$ 項 $\Rightarrow$ $Q_2=6$

$\boldsymbol{M}$e在第 $\dfrac{2\times 7}{4}+\dfrac{1}{2}=4$ 項 $\Rightarrow$ $Q_2=6$

$\boldsymbol{D_5}$在第 $\dfrac{5\times 7}{10}+\dfrac{1}{2}=4$ 項 $\Rightarrow$ $Q_5=6$

$\boldsymbol{P_{50}}$在第 $\dfrac{50\times 7}{100}+\dfrac{1}{2}=4$ 項 $\Rightarrow$ $P_{50}=6$

$\therefore\quad Q_2=M_e=D_5=P_{50}$

3-7 眾數

- **眾數的意義及計算方式：**

眾數 (mode)，係指一群資料中出現最多次的數值(即高峰所在)，稱之。

若二相鄰兩數值均為出現次數最多的數值，則取其算術平均數為其眾數 (M_o)。

例題 3-14

(a) 試求1, 3, 3, 7, 7, 7, 4, 4, 8之眾數。

(b) 試求1, 3, 3, 7, 7, 8, 10之眾數。

(c) 試求2, 5, 5, 5, 7, 8, 8, 8, 9之眾數。

(d) 試求3, 4, 7, 6, 5之眾數。

解：(a) $M_o = 7$ (有一眾數，即單峰分配)

(b) $M_o = 5$ (因3與7為相鄰二數值，故取其算術平均數 $\dfrac{3+7}{2} = 5$ 為眾數)

(c) $M_o = 5$ 和8 (有二眾數，即雙峰分配)

(d) 沒有眾數 (該分配沒有高峰)

3-8 分組資料平均數之計算

- 分組資料算術平均數的計算方式：
  - 母體資料：

$$\mu = \frac{1}{N}\sum_{i=1}^{K} f_i X_i = \frac{1}{N}\sum f X$$

  - 樣本資料：

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{K} f_i X_i = \frac{1}{n}\sum f X$$

其中　X_i = 第 i 組的組中點

N (或n) = $\sum f$ = 總次數

f_i = 第 i 組的次數

K = 組數

例題 3-15

甲班50名學生統計學成績如下表所示，試求算術平均數。

成績(分)	人數
25~35	2
35~45	10
45~55	18
55~65	16
65~75	3
75~85	1
合　計	50

解：

成績（分）	人數f	組中點X	fX
25~35	2	30	60
35~45	10	40	400
45~55	18	50	900
55~65	16	60	960
65~75	3	70	210
75~85	1	80	80
合　計	50		2,610

$$\mu = \frac{1}{N}\sum f\ X = \frac{1}{50} \times 2{,}610 = 52.2\ (\text{分})$$

3-8 分組資料平均數之計算

- **分組資料分位數的計算方式：**

$$P_r = L_{pr} + \frac{\frac{rn}{100} - F_{pr}}{f_{pr}} \times h_{pr}$$

其中 P_r = 第 r 個百分位數

L_{pr} = P_r 所在組的下限

F_{pr} = 組中點小於 L_{pr} 之各組次數之總和

f_{pr} = P_r 所在組之次數

h_{pr} = P_r 所在組之組距

- **分組資料眾數的計算方式：**

$$M_o = \mu - 3(\mu - M_e)$$

例題 3-16

試就上例求算 Q_1 , M_e , Q_3 , D_5 , P_{25} , P_{50} , P_{75} 。

解：

成績 (分)	人數f	以下累積
25~35	2	2
35~45	10	12
45~55	18	30
55~65	16	46
65~75	3	49
75~85	1	50
合　計	50	

(a) 公式解：

① 先確定 $P_{25}(=Q_1)$, $P_{50}(=M_e=D_5)$, $P_{75}(=Q_3)$ 在那一組內

P_{25} 在第 $\dfrac{25\times 50}{100}=12.5$ 項，屬第3組

P_{50} 在第 $\dfrac{50\times 50}{100}=25$ 項，屬第3組

P_{75} 在第 $\frac{75\times 50}{100}=37.5$ 項，屬第4組

② 代入公式

$$P_{25}=L_{P25}+\frac{\frac{25n}{100}-F_{P25}}{f_{P25}}\times h_{P25}=45+\frac{\frac{25\times 50}{100}-12}{18}\times 10$$

$$=45.28\ (\text{分})$$

$$P_{50}=L_{P50}+\frac{\frac{50n}{100}-F_{P50}}{f_{P50}}\times h_{P50}=45+\frac{\frac{50\times 50}{100}-12}{18}\times 10$$

$$=52.22\ (\text{分})$$

$$P_{75}=L_{P75}+\frac{\frac{75n}{100}-F_{P75}}{f_{P75}}\times h_{P75}=55+\frac{\frac{75\times 50}{100}-30}{16}\times 10$$

$$=59.69\ (\text{分})$$

(b) 插補法：

$$P_{25} \Rightarrow \begin{pmatrix} 45 & 12 \\ P_{25} & 12.5 \\ 55 & 30 \end{pmatrix} \xRightarrow{\text{插補法}} \frac{P_{25}-45}{55-45}=\frac{12.5-12}{30-12}$$

(組距)(累積次數)

$$\therefore \quad P_{25}=45+\frac{12.5-12}{30-12}\times(55-45)$$

$$=45+\frac{12.5-12}{18}\times 10=45.28\text{ (分)}$$

$$P_{50} \Rightarrow \begin{pmatrix} 45 & 12 \\ P_{50} & 25 \\ 55 & 30 \end{pmatrix} \xRightarrow{\text{插補法}} \frac{P_{50}-45}{55-45}=\frac{25-12}{30-12}$$

$$\therefore \quad P_{50}=45+\frac{25-12}{30-12}\times(55-45)$$

$$=45+\frac{25-12}{18}\times 10=52.22\text{ (分)}$$

$$P_{75} \Rightarrow \begin{pmatrix} 55 & 30 \\ P_{75} & 37.5 \\ 65 & 46 \end{pmatrix} \xRightarrow{\text{插補法}} \frac{P_{75}-55}{65-55}=\frac{37.5-30}{46-30}$$

$$\therefore \quad P_{75} = 55 + \frac{37.5-30}{46-30} \times (65-55)$$

$$= 55 + \frac{37.5-30}{16} \times 10 = 59.69\ (\text{分})$$

(c) 由公式解與插補法計算可知，兩者相同。在前面我們曾說明分組資料計算分位數是假設每組內的資料是均勻分佈的，所以計算方式可由插補法求算，這也是公式解的由來。

例題 3-17

假設根據某統計資料計算出，試以皮爾生法求算眾數，並判斷次數分配形態。

解：(a) $M_o = \mu - 3(\mu - M_e) = 60 - 3(60 - 58.5) = 55.5$

(b) $\because \mu > M_e > M_o$ $\therefore$是單峰右偏次數分配。[詳見以下3-9節之說明]

3-9 各種平均數之關係與比較

枝葉圖 (stem-and-leaf plot)

又稱為莖葉圖，
它是一種結合數字與圖形的統計資料之表達圖示。
其外觀與橫式直方圖類似，
但比橫式直方圖能夠提供更多統計資料的特性。

例題 3-18

試以箱形圖型式表達下列一群統計資料：

34，48，7，37，100，51，65，49

解： 表達箱形圖首先要求算四分位數。

(a) 先將資料從小到大排列

7, 34, 37, 48, 49, 51, 65, 100

(b) ① Q_1 在第 $\frac{8}{4}+\frac{1}{2}=2.5$ 項

$$Q_1=34+0.5(37-34)=35.5$$

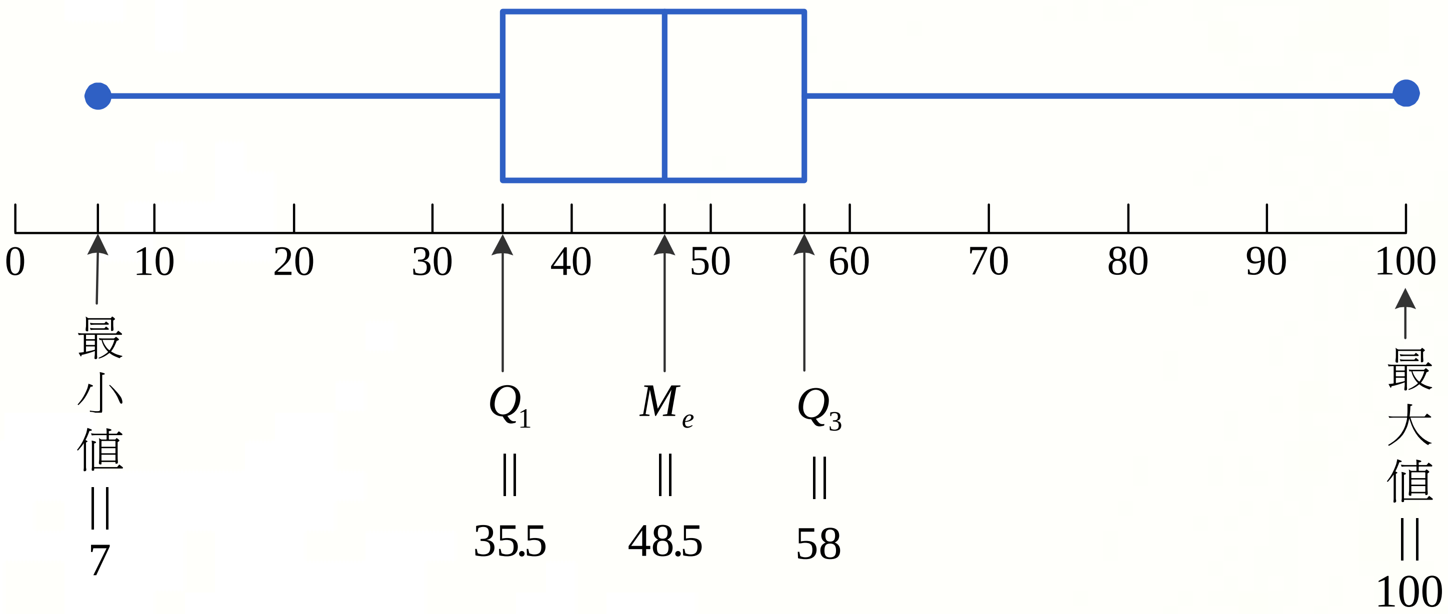
② $Q_2=M_e$ 在第 $\frac{2\times 8}{4}+\frac{1}{2}=4.5$ 項

$$Q_2=48+0.5(49-48)=48.5$$

③ Q_3 在第 $\frac{3\times 8}{4}+\frac{1}{2}=6.5$ 項

$$Q_3=51+0.5(65-51)=58$$

(c) 箱形圖表達如下圖所示：



由上圖可看出資料中7與100離 Q_1 和 Q_3 太遠，有極端值出現的情形。