B2-3-1

一維數據分析【基礎篇】

【主題 1】平均數、中位數與百分位數

焦點 1. 母體與樣本

一般來說，統計問題的數據分為【母體數據】及【樣本數據】。母體是指我們所欲尋求資訊的標的，當母體很大、難以掌握時，我們可以用適當的抽樣方式從其中抽出一部分來檢視，叫做樣本，然後想辦法從樣本中推論出關於母體的資訊。本章所考慮的數據，若未特別聲明，均視為母體數據。

焦點 2. 算術平均數、中位數

最常用來當作數據「代表值」的數字，有【算術平均數】和【中位數】兩種。

1. 算術平均數（簡稱為平均數）

(1) 假設數值資料 X 中包含 n 個數據：$x_1, x_2, \cdots, x_n$，則其算術平均數為

$$\frac{x_1+x_2+\cdots+x_n}{n}。$$

註：通常樣本平均數以符號 $\overline{x}$（唸做 x bar）表示，
母體平均數則以希臘字母 μ（音似 mu）來表示。

(2) 算術平均數容易受到少數【極端值】的影響。

3. 加權平均數

設有 n 個數據 x_1，x_2，⋯，x_n，其對應的權數分別為 w_1，w_2，⋯，w_n，則加權平均數為

$$W=\frac{x_1w_1+x_2w_2+\cdots+x_nw_n}{w_1+w_2+\cdots+w_n}。$$

4. 幾何平均數

設有 n 個數據 x_1，x_2，⋯，x_n，且都是正數，其幾何平均數為 $G=\sqrt[n]{x_1x_2\cdots x_n}$

註：① 我們常利用幾何平均數來計算平均成長率。

② 若某物件一開始的價值為 p，連續 n 年的成長率分別為 r_1，r_2，⋯，r_n，假設每年的平均成長率為 r，則有 $p(1+r)^n=p(1+r_1)(1+r_2)\cdots(1+r_n)$

$\therefore\ r=\sqrt[n]{(1+r_1)(1+r_2)\cdots(1+r_n)}-1$

2. 中位數

假設數值資料 X 中包含 n 個數據：$x_1, x_2, \cdots, x_n$，若此 n 個數據從小到大排序之後為 $x_{(1)}, x_{(2)}, \cdots, x_{(n)}$，則其中位數

$$\mathrm{Me}=\begin{cases} x_{(\frac{n+1}{2})} & ,n\text{ 為奇數} \\ \dfrac{x_{(\frac{n}{2})}+x_{(\frac{n}{2}+1)}}{2} & ,n\text{ 為偶數}\end{cases}。$$

註：平均數容易受到少數【極端值】的影響，中位數則較不受影響。因此，為了避免過大或過小的極端值影響我們對數據「代表值」的估計，有時候會採取中位數而非平均數。

【範例 1】算術平均數、中位數

有一組數據如下：
2563, 2447, 2512, 2400, 3000, 2504 試求此組數據的
(1) 算術平均數。　(2) 中位數。

(1) 所求 $=\dfrac{2563+2447+2512+2400+3000+2504}{6}=2571$

(2) 排序：2400, 2447, 2504, 2512, 2563, 3000
$\because n=6$ 是偶數，$\therefore$ 中位數等於排序後第 3 個數和第 4 個數的平均 $=\dfrac{2504+2512}{2}=2508$

【範例 2】算術平均數、中位數

已知 5 個數據如下：30, 25, 15, 40 及 x，若此 5 個數的算術平均數與中位數相等，則 x 的可能值為何？

解題關鍵：依 x 的範圍分段討論：
(1) $x<25$ 或 (2) $25\le x\le 30$ 或 (3) $x>30$

此 5 個數的算術平均數
$$\mu=\frac{30+25+15+40+x}{5}=\frac{110+x}{5}=22+\frac{x}{5}$$
先將已知的 4 個數由小到大排序之後得 15, 25, 30, 40

(1) 若 $x<25$，則 $\mathrm{Me}=25=\mu=22+\dfrac{x}{5}\Rightarrow\dfrac{x}{5}=3\Rightarrow x=15$

(2) 若 $25\le x\le 30$，則 $\mathrm{Me}=x=\mu=22+\dfrac{x}{5}\Rightarrow\dfrac{4x}{5}=22\Rightarrow x=27.5$

(3) 若 $x>30$，則 $\mathrm{Me}=30=\mu=22+\dfrac{x}{5}\Rightarrow\dfrac{x}{5}=8\Rightarrow x=40$

由(1), (2), (3)可知，x 的可能值為 15 或 27.5 或 40

【範例 3】加權平均數

某次月考某生成績及每週上課時數如下表，若該生成績的加權平均數介於 60~65 分之間，則數學成績可能是幾分？(多選)

科目	國文	英文	數學	自然	社會
成績	70	50	x	85	50
時數	6	6	5	4	4

(A) 45 分　(B) 49 分　(C) 55 分　(D) 70 分

$$W=\frac{6\times70+6\times50+5\times x+4\times85+4\times50}{6+6+5+4+4}=\frac{1260+5x}{25}$$

$$\Rightarrow\quad 60<\frac{1260+5x}{25}<65\quad\Rightarrow\quad 1500<1260+5x<1625$$

$\Rightarrow\quad 48<x<73$　　故選(B)(C)(D)

【範例 4】幾何平均數(平均成長率)

某件物品連續四年的售價為 1000、1200、1500、1728 元，則該物品
(1) 平均每年的變動率為______%。
(2) 每年物品價格平均增加______%。

(1) $G=\sqrt[3]{\frac{1200}{1000}\times\frac{1500}{1200}\times\frac{1728}{1500}}=\sqrt[3]{\frac{1728}{1000}}=1.2=120\%$

(2) $r=G-1=1.2-1=0.2=20\%$

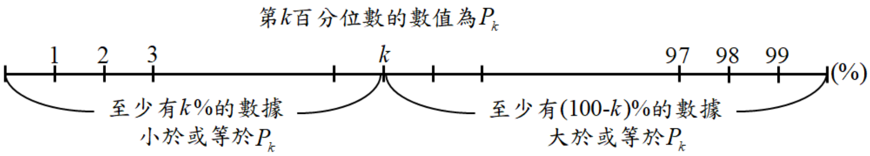
$$r=\sqrt[3]{(1+\frac{1200-1000}{1000})(1+\frac{1500-1200}{1200})(1+\frac{1728-1500}{1500})}-1$$

$$=\sqrt[3]{(\frac{1200}{1000})(\frac{1500}{1200})(\frac{1728}{1500})}-1=1.2-1=0.2=20\%$$

焦點 3. 百分位數與四分位數

1. 百分位數

(1) 若 P_k 是第 k 百分位數（$k=1,2,\cdots,99$），則至少有 $k\%$ 的數值資料小於或等於 P_k，同時也至少有 $(100-k)\%$ 的數值資料大於或等於 P_k。



(2) 假設數值資料 X 中包含 n 個數據：x_1, x_2, $\cdots$, x_n，若此 n 個數據從小到大排序之後為 $x_{(1)}$, $x_{(2)}$, $\cdots$, $x_{(n)}$，令 P_k 代表第 k 百分位數，且 $t=n\times\frac{k}{100}$，則

$$P_k=\begin{cases}x_{([t]+1)} & ,t\text{ 不是整數}\\ \dfrac{x_{(t)}+x_{(t+1)}}{2} & ,t\text{ 是整數}\end{cases}$$ 。

註：上式中的 $[t]$ 稱為 t 取高斯符號，表示不大於 t 的最大整數。
　$\boxed{\text{例如}}$ $[3.7]=3$，$[4]=4$，$[-2.3]=-3$。

2. 四分位數

假設數值資料 X 中包含 n 個數據，從小到大排序後為 $x_{(1)}$, $x_{(2)}$, $\cdots$, $x_{(n)}$，則
(1) 第 25 百分位數 P_{25} 亦稱為第 1 四分位數，記為 Q_1。
(2) 第 50 百分位數 P_{50} 即中位數 Me，亦稱為第 2 四分位數，記為 Q_2。
(3) 第 75 百分位數 P_{75} 亦稱為第 3 四分位數，記為 Q_3。

【範例 5】百分位數

假設數值資料 X 中包含 150 個數據，其排序後如下：

$x_{(1)}=1$，$x_{(2)}=2$，$x_{(3)}=3$，$x_{(4)}=4$，$x_{(5)}=5$，…，$x_{(k)}=k$，…，$x_{(150)}=150$，試求：

(1) 第 3 百分位數 P_3。

(2) 第 10 百分位數 P_{10}。

(1) $\because t=150\times\dfrac{3}{100}=4.5$ 不是整數

$\therefore P_3=x_{([4.5]+1)}=x_{(5)}=5$

(2) $\because t=150\times\dfrac{10}{100}=15$ 為整數

$\therefore P_{10}=\dfrac{x_{(15)}+x_{(16)}}{2}=\dfrac{15+16}{2}=15.5$

【範例 6】平均數、中位數與四分位數

調查 50 個家庭的子女數，其統計表如下：

子女數（人）	0	1	2	3	4	5	6
次數（家）	3	8	28	5	3	2	1

試求子女數的：

(1) 算術平均數。　(2) 中位數。

(3) 第 1 四分位數 Q_1 與第 3 四分位數 Q_3。

(1) $\dfrac{0\times3+1\times8+2\times28+3\times5+4\times3+5\times2+6\times1}{50}=2.14$（人）

(2) 中位數等於排序後第 25 個數 $x_{(25)}$ 和第 26 個數 $x_{(26)}$ 的平均 $=\dfrac{2+2}{2}=2$（人）

(3) $\because 50\times\dfrac{25}{100}=50\times\dfrac{1}{4}=12.5$ 不是整數

$\therefore Q_1=P_{25}=x_{([50\times\frac{1}{4}]+1)}=x_{(13)}=2$（人）

，$Q_3=P_{75}=x_{([50\times\frac{3}{4}]+1)}=x_{(38)}=2$（人）

NOTE

【主題 2】母體變異數與標準差

(1) 假設一群數值 x_1， x_2， $\cdots$， x_n 的算術平均數為 μ ，

則 $x_i-\mu$ 稱為 x_i （ $i=1,2,\cdots,n$ ）的離均差(離差)。

(2) 離均差(離差)總和為 0。即

$$(x_1-\mu)+(x_2-\mu)+\cdots+(x_n-\mu)=(x_1+x_2+\cdots+x_n)-n\mu=0 \text{ 。}$$

(3) 離均差(離差)的平方和，不大於任何其他數值與變量差之平方和

$$(x_1-\mu)^2+(x_2-\mu)^2+\cdots+(x_n-\mu)^2\le(x_1-A)^2+(x_2-A)^2+\cdots+(x_n-A)^2$$

焦點 2. 母體變異數與標準差公式

假設母體數據 x_1， x_2， $\cdots$， x_n 的平均數為 μ ，則：

(1) 母體變異數 $\sigma^2=\dfrac{(x_1-\mu)^2+(x_2-\mu)^2+\cdots+(x_n-\mu)^2}{n}$

$$=\frac{{x_1}^2+{x_2}^2+\cdots+{x_n}^2}{n}-\mu^2 \text{ 。}$$

(2) 標準差 $\sigma=\sqrt{\dfrac{(x_1-\mu)^2+(x_2-\mu)^2+\cdots+(x_n-\mu)^2}{n}}$

$$=\sqrt{\frac{{x_1}^2+{x_2}^2+\cdots+{x_n}^2}{n}-\mu^2} \text{ 。}$$

註 1： σ 近似英文 sigma 的發音。

註 2：當標準差較大時，代表每個數據離平均數相對較遠，所以數據資料較分散；當標準差較小時，代表每個數據離平均數相對較近，所以數據資料較集中。反之亦然。

【範例 1】母體變異數與標準差

試求下列數據的母體變異數與標準差： 25, 30 , 15, 40 , 40。

平均數 $\mu=\dfrac{25+30+15+40+40}{5}=30$

母體變異數

$$\sigma^2=\frac{1}{5}[(25-30)^2+(30-30)^2+(15-30)^2+(40-30)^2+(40-30)^2]$$

$$=\frac{450}{5}=90$$

標準差 $\sigma=\sqrt{90}=3\sqrt{10}$

NOTE

【範例 2】母體變異數與標準差

下列 5 組資料（每組各有 10 筆），試問哪一組資料的母體標準差最大？
(A) 1, 1, 1, 1, 1, 10, 10, 10, 10, 10
(B) 1, 1, 1, 1, 1, 5, 5, 5, 5, 5
(C) 4, 4, 4, 5, 5, 5, 5, 6, 6, 6
(D) 1, 1, 2, 2, 3, 3, 4, 4, 5, 5
(E) 1, 2, 3, 4, 5, 6, 7, 8, 9, 10　　　　　　　　　　【學測】

比較各選項中 10 個資料的分散程度
∵(A)選項中 10 個資料的分散程度最大，
∴標準差最大
故選(A)

【範例 3】母體標準差的計算公式

(1) 設一組母體資料有 50 個數據，其總和為 150，平方和為 6500，試求此組資料的標準差為何？
(2) 已知某班共有 40 人，某次數學考試平均為 65 分，標準差為 13 分，試問其分數之平方和是多少？

(1) 設此組資料的 50 個數據為 x_1，x_2，$\cdots$，x_{50}

則其平均數 $\mu=\dfrac{x_1+x_2+\cdots+x_{50}}{50}=\dfrac{150}{50}=3$

母體變異數 $\sigma^2=\dfrac{{x_1}^2+{x_2}^2+\cdots+{x_{50}}^2}{50}-\mu^2$

$=\dfrac{6500}{50}-3^2=121$

$\Rightarrow$ 標準差 $\sigma=\sqrt{121}=11$

(2) 設全班 40 人的分數分別為 x_1，x_2，$\cdots$，x_{40}

則母體變異數 $\sigma^2=\dfrac{{x_1}^2+{x_2}^2+\cdots+{x_{40}}^2}{40}-\mu^2$

$\Rightarrow 13^2=\dfrac{{x_1}^2+{x_2}^2+\cdots+{x_{40}}^2}{40}-65^2$

$\Rightarrow {x_1}^2+{x_2}^2+\cdots+{x_{40}}^2=40\times(13^2+65^2)$

$=175760$（分^2）

NOTE

【範例 4】分組合併標準差之計算

某次測驗分成甲、乙二組，甲組有 30 人，平均成績為 75 分，標準差為 4 分；乙組有 20 人，平均成績為 80 分，標準差為 1 分。試求此次測驗 50 人的
(1) 平均成績。 (2) 母體標準差。

解題關鍵：先求出各組數據的平方和，將各組數據平方和相加，
便可求出合併後全部數據的平方和。

設甲組 30 人的成績分別為 $x_1,\ x_2,\ \cdots,\ x_{30}$，
平均成績為 $\mu_{甲}$，標準差為 $\sigma_{甲}$

乙組 20 人的成績分別為 $x_{31},\ x_{32},\ \cdots,\ x_{50}$，
平均成績為 $\mu_{乙}$，標準差為 $\sigma_{乙}$

(1) 此 50 人的平均成績

$$\mu_{合}=\frac{30\times\mu_{甲}+20\times\mu_{乙}}{50}=\frac{30\times75+20\times80}{50}=77\text{（分）}$$

(2) $\because\ 4^2=\sigma_{甲}{}^2=\dfrac{x_1{}^2+x_2{}^2+\cdots+x_{30}{}^2}{30}-75^2$

$\therefore\ x_1{}^2+x_2{}^2+\cdots+x_{30}{}^2=30\times(75^2+4^2)=30\times5641$

又 $1^2=\sigma_{乙}{}^2=\dfrac{x_{31}{}^2+x_{32}{}^2+\cdots+x_{50}{}^2}{20}-80^2$

$\therefore\ x_{31}{}^2+x_{32}{}^2+\cdots+x_{50}{}^2=20\times(80^2+1^2)=20\times6401$

因此，此 50 人的母體變異數 $\sigma_{合}{}^2=\dfrac{x_1{}^2+x_2{}^2+\cdots+x_{50}{}^2}{50}-\mu_{合}{}^2$

$$=\frac{(x_1{}^2+\cdots+x_{30}{}^2)+(x_{31}{}^2+\cdots+x_{50}{}^2)}{50}-77^2=\frac{30\times5641+20\times6401}{50}-77^2=16\text{（分}^2\text{）}$$

$\Rightarrow$ 標準差 $\sigma=\sqrt{16}=4$（分）

NOTE

【主題 3】線性變換與數據的標準化

焦點 1. 數據的平移與伸縮

已知一組母體數據為 x_1, x_2, $\cdots$, x_n，若將每一筆資料同時加（或減）一個定數稱為平移，而同時乘以一個非零的常數則稱為伸縮。

焦點 2. 數據的線性變換

(1) 已知一組母體數據為 x_1, x_2, $\cdots$, x_n，
若將每一筆資料同時乘以一個常數 a 再加上一個常數 b
而成另一組資料 $y_i = ax_i + b$（$i = 1, 2, \cdots, n$），
則這種轉換稱為線性變換。

(2) 假設有母體數據 x_1, x_2, $\cdots$, x_n，其平均數為 μ_x、變異數為 σ_x^2、標準差為 σ_x；
令 $y_i = ax_i + b$（$i = 1, 2, \cdots, n$），而 y_1, y_2, $\cdots$, y_n 的
平均數為 μ_y、變異數為 σ_y^2、標準差為 σ_y，則：
① $\mu_y = a\mu_x + b$。
② ${\sigma_y}^2 = a^2{\sigma_x}^2$。
③ $\sigma_y = |a|\sigma_x$。

焦點 3. 數據標準化與標準分數

假設有母體數據 x_1, x_2, $\cdots$, x_n，其平均數為 μ_x、標準差為 σ_x。

(1) 將數據標準化，指的是將數據減去平均數之後，再除以標準差，即

$$\frac{x_i - \mu_x}{\sigma_x} \quad (i = 1, 2, \cdots, n)$$

(2) 數據 x_i 的標準分數為 $z_i = \dfrac{x_i - \mu_x}{\sigma_x}$。

(3) 標準分數的性質：標準分數 z_1, z_2, $\cdots$, z_n 的
① 算術平均數 $\mu_z = 0$。
② 標準差 $\sigma_z = 1$。

(4) 標準分數的意義
從標準分數的計算方式便可看出，標準分數是以整組數據的平均數當作「原點」，而以標準差當作度量單位的一種表達方式。
例如：標準分數的平均數必定等於 0；當一個數的標準分數等於 1 時，代表這個數比平均數大 1 個標準差，而當一個數的標準分數等於 -2 時，代表這個數比平均數小 2 個標準差。
標準分數可以用來比較兩個分別屬於不同母體的數據。

【範例 1】數據的線性變換

一組母體資料有 n 個數據 x_1，x_2，$\cdots$，x_n，其算術平均數為 40，中位數為 45，標準差為 3。試求以下新數據之算術平均數、中位數及標準差：

(1) $2x_1, 2x_2, \cdots, 2x_n$。

(2) x_1+4，x_2+4，$\cdots$，x_n+4。

(3) $\frac{1}{2}x_1+1$，$\frac{1}{2}x_2+1$，$\cdots$，$\frac{1}{2}x_n+1$。

(1) 令 $y_i=2x_i$（$i=1,2,\cdots,n$），則新數據的

平均數 $\mu_y=2\mu_x=2\times 40=80$

中位數 $\mathrm{Me}_y=2\,\mathrm{Me}_x=2\times 45=90$

標準差 $\sigma_y=|2|\sigma_x=2\times 3=6$

(2) 令 $y_i=x_i+4$（$i=1,2,\cdots,n$），則新數據的

平均數 $\mu_y=\mu_x+4=40+4=44$

中位數 $\mathrm{Me}_y=\mathrm{Me}_x+4=45+4=49$

標準差 $\sigma_y=\sigma_x=3$

(3) 令 $y_i=\frac{1}{2}x_i+1$（$i=1,2,\cdots,n$），則新數據的

平均數 $\mu_y=\frac{1}{2}\mu_x+1=\frac{1}{2}\times 40+1=21$

中位數 $\mathrm{Me}_y=\frac{1}{2}\,\mathrm{Me}_x+1=\frac{1}{2}\times 45+1=\frac{47}{2}$

標準差 $\sigma_y=|\frac{1}{2}|\sigma_x=\frac{1}{2}\times 3=\frac{3}{2}$

【範例 2】數據標準化

已知一組數據 1, 3, 5, 7, 4，試求此組數據的標準化數據。

此組數據的

平均數 $\mu=\dfrac{1+3+5+7+4}{5}=4$

變異數 $\sigma^2=\dfrac{(-3)^2+(-1)^2+1^2+3^2+0^2}{5}=4$

標準差 $\sigma=\sqrt{4}=2$

因此，將此組數據標準化得

$\frac{1-4}{2}$，$\frac{3-4}{2}$，$\frac{5-4}{2}$，$\frac{7-4}{2}$，$\frac{4-4}{2}$

即 -1.5，$-0.5, 0.5, 1.5, 0$

【範例 3】標準分數

某校補救教學班包含阿三共有 6 位同學參加。某次總分 100 分的小考從最高到最低分如下：86, 86, 80, 80, 80, 68。若阿三的成績是其中的 86 分，則他的標準分數為何？

此 6 位同學分數的

平均數 $\mu=\dfrac{86+86+80+80+80+68}{6}=80$（分）

變異數 $\sigma^2=\dfrac{6^2+6^2+0^2+0^2+0^2+(-12)^2}{6}=36$（分^2）

標準差 $\sigma=\sqrt{36}=6$（分）

因此，阿三的標準分數為 $\dfrac{86-80}{6}=1$

【範例 4】標準分數

某班數學成績平均為 65 分，標準差為 15 分；英文成績平均為 72 分，標準差為 11 分。已知阿三的數學成績為 80 分，英文成績為 80 分。

(1) 試將阿三的這兩科成績標準化。

(2) 阿三哪一科的表現較好？

數學成績的標準分數為 $\dfrac{80-65}{15}=1$

英文成績的標準分數為 $\dfrac{80-72}{11}=\dfrac{8}{11}$

$\because$ 阿三的數學成績高於全班平均 1 個標準差，

英文成績高於全班平均 $\dfrac{8}{11}$ 個標準差

$\therefore$ 數學成績的表現較好

NOTE

【主題 4】函數的最大、最小值問題

(1) 設函數 $f(x)=(x-x_1)^2+(x-x_2)^2+\cdots+(x-x_n)^2$，當

$x=\dfrac{x_1+x_2+\cdots+x_n}{n}$（即 x_1，x_2，$\cdots$，x_n 的算術平均數）時，$f(x)$ 有最小值。

(2) 設 $x_1\le x_2\le\cdots\le x_n$，$f(x)=|x-x_1|+|x-x_2|+\cdots+|x-x_n|$。

① 若 n 為奇數，則當 $x=x_{\frac{n+1}{2}}$（即 x_1，x_2，$\cdots$，x_n 的中位數）時，$f(x)$ 有最小值。

② 若 n 為偶數，則當 $x_{\frac{n}{2}}\le x\le x_{\frac{n}{2}+1}$（$x_1$，$x_2$，$\cdots$，$x_n$ 的中位數亦在此範圍內）時，$f(x)$ 有最小值。

【範例 1】二次函數的最小值

設函數 $f(x)=(x-1.1)^2+(x-2.2)^2+(x-3.3)^2+(x-4.4)^2+(x-5.5)^2$，
當 $f(x)$ 有最小值時，試求此時 x 的值。

當 $x=\dfrac{\overbrace{1.1+2.2+3.3+4.4+5.5}^{\text{等差級數，公差=1.1}}}{5}=\dfrac{\frac{5\times(1.1+5.5)}{2}}{5}=3.3$ 時，$f(x)$ 有最小值

【範例 2】二次函數的最小值

設函數 $f(x)=1\cdot(x-1)^2+2\cdot(x-2)^2+\cdots+10\cdot(x-10)^2$，當 $f(x)$ 有最小值時，試求此時 x 的值。

$$
\begin{aligned}
f(x)&=1\cdot(x-1)^2+2\cdot(x-2)^2+\cdots+10\cdot(x-10)^2\\
&=(x-1)^2+\underbrace{(x-2)^2+(x-2)^2}_{2\text{個}}+\cdots+\underbrace{(x-10)^2+\cdots+(x-10)^2}_{10\text{個}}
\end{aligned}
$$

因此，當
$$
\begin{aligned}
x&=\frac{1+\overbrace{2+2}^{2\text{個}}+\overbrace{3+3+3}^{3\text{個}}+\cdots+\overbrace{10+\cdots+10}^{10\text{個}}}{1+2+\cdots+10}\\
&=\frac{1\times1+2\times2+3\times3+\cdots+10\times10}{\frac{10\times(1+10)}{2}}\\
&=\frac{1^2+2^2+3^2+\cdots+10^2}{55}\\
&=\frac{\frac{10\times(10+1)\times(2\times10+1)}{6}}{55}=7
\end{aligned}
$$
時 $f(x)$ 有最小值

【範例 3】絕對值函數的最小值

(1) 設函數 $f(x)=|x-1|+|x-2|+|x-5|$，試求 $f(x)$ 的最小值，以及此時 x 的值。

(2) 若函數 $f(x)=|x-1|+|x-2|+|x-5|+|x-8|$，試求 $f(x)$ 的最小值，以及此時 x 的值。。

(1) $\because 1, 2, 5$ 三個數的中位數為 2

$\therefore$ 當 $x=2$ 時，$f(x)$ 有最小值　$f(2)=|2-1|+|2-2|+|2-5|=4$

(2) $\because 1, 2, 5, 8$ 有四個數

$\therefore$ 當 $2\le x\le 5$ 時，$f(x)$ 有最小值 $f(2)$（或 $f(5)$）$=|2-1|+|2-2|+|2-5|+|2-8|$
$=10$

【範例 4】絕對值函數的最小值

若函數 $f(x)=|2x+1|+|3x-2|+|4x-5|$，試求 $f(x)$ 的最小值，以及此時 x 的值。。

$$f(x)=|2x+1|+|3x-2|+|4x-5|$$

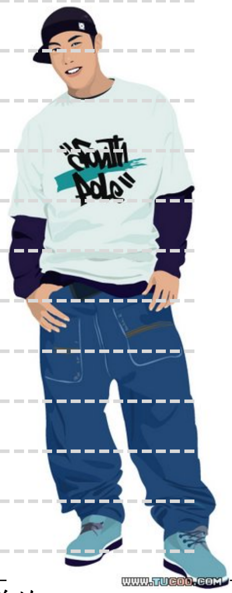
$$=2|x+\frac{1}{2}|+3|x-\frac{2}{3}|+4|x-\frac{5}{4}|$$

$$=\underbrace{\left|x+\frac{1}{2}\right|+\left|x+\frac{1}{2}\right|}_{2個}+\underbrace{|x-\frac{2}{3}|+\cdots+|x-\frac{2}{3}|}_{3個}+\underbrace{|x-\frac{5}{4}|+\cdots+|x-\frac{5}{4}|}_{4個}$$

$\because -\frac{1}{2},-\frac{1}{2},\frac{2}{3},\frac{2}{3},\frac{2}{3},\frac{5}{4},\frac{5}{4},\frac{5}{4},\frac{5}{4}$ 九個數的中位數為 $\frac{2}{3}$

$\therefore$ 當 $x=\frac{2}{3}$ 時，$f(x)$ 有最小值 $=f(\frac{2}{3})=\frac{14}{3}$

NOTE



3-1 自我評量

[基礎題]

1.某班共有 50 位同學，全班平均身高是 168 公分，但有兩位同學轉出，身高分別是 165 公分與 166 公分，後來又轉進兩位新同學，其身高分別為 169 公分與 x 公分，結果全班平均身高變為 168.2 公分，則 $x=$________。

解

$$\frac{50\times168-165-166+169+x}{50}=168.2$$

得 $x=172$

2.________若已知某一筆數據之算術平均數 $\mu_x=10$，標準差 $S_x=3$，中位數 $Me_x=12$ 第一四分位數 $Q_{1x}=8$，若 $y=-4x+3$。則對新數據 Y 而言，下列選項何者正確？

(A)算術平均數 $\mu_y=43$　(B)標準差 $\sigma_y=-12$

(C)中位數 $Me_y=-45$　(D)第一四分位數 $Q_{1y}=-32$

解

(A)×：$\mu_y=-4\mu_x+3=-37$

(B)×：$\sigma_y=|-4|\sigma_x=12$

(C)○：$Me_y=-4Me_x+3=-45$

(D)×：$Q_{1y}=-4Q_{x1}+3=-29$

故選(C)

3.某甲第一次段考六科成績的算術平均為 80，若已知其中有五科成績為 68, 80, 80, 80, 86，則第六科的成績為________，又成績的標準差為________。

解

設第六科成績為 x

平均數 $\frac{68+80+80+80+86+x}{6}=80\Rightarrow x=86$

標準差 $\sigma=\sqrt{\frac{(-12)^2+0^2+0^2+0^2+6^2+6^2}{6}}=6$

4.________國中會考有 30 萬名學生應考，若小敏的成績是超過第 86 百分位數，則小敏在這 30 萬名考生中至少贏過幾人？

(A)39000　(B)42000　(C)258000　(D)261000

解

$300000\times\frac{86}{100}=258000$，可知至少贏過 258000

故選(C)

5.__________根據臺灣壽險業的資料，男性從 0 歲、1 歲、…到 60 歲各年齡層的死亡率（單位：%）依序為

1.0250，0.2350，0.1520，0.1010，0.0720，0.0590，0.0550，0.0540，0.0540，0.0520，0.0490，0.0470，0.0490，0.0560，0.0759，0.1029，0.1394，0.1890，0.2034，0.2123，0.2164，0.2166，0.2137，0.2085，0.2019，0.1948，0.1882，0.1830，0.1799，0.1793，0.1813，0.1862，0.1941，0.2051，0.2190，0.2354，0.2539，0.2742，0.2961，0.3202，0.3472，0.3779，0.4129，0.4527，0.4962，0.5420，0.5886，0.6346，0.6791，0.7239，0.7711，0.8229，0.8817，0.9493，1.0268，1.1148，1.2139，1.3250，1.4485，1.5851，1.7353。

經初步整理後，已知 61 個資料中共有 24 個資料小於 0.2。請問死亡率資料的中位數為下列哪一個選項？

(A)0.2034　(B)0.2164　(C)0.2137　(D)0.2085　(E)0.2019

解

$61\times\frac{50}{100}=30.5\Rightarrow Me=x_{([30.5]+1)}=x_{(31)}$，所以要找小到大第 31 個資料

但小於 0.2 的資料有 24 個，所以由小到大超過 0.2 的第 7 筆資料，可得 $Me=0.2164$

故選(B)

6.___________根據一百多年來的氣象紀錄，美國費城年雨量平均值為 41.0 英吋，標準差為 6.1 英吋。今欲將此項統計資料的單位由英制換為公制，試問該城市一百多年來年雨量的標準差最接近下列的哪一個選項?

（註：1 英吋等於 25.4 毫米）

(A)0.240 毫米　(B)1.61 毫米　(C)6.10 毫米　(D)155 毫米　(E)1041 毫米

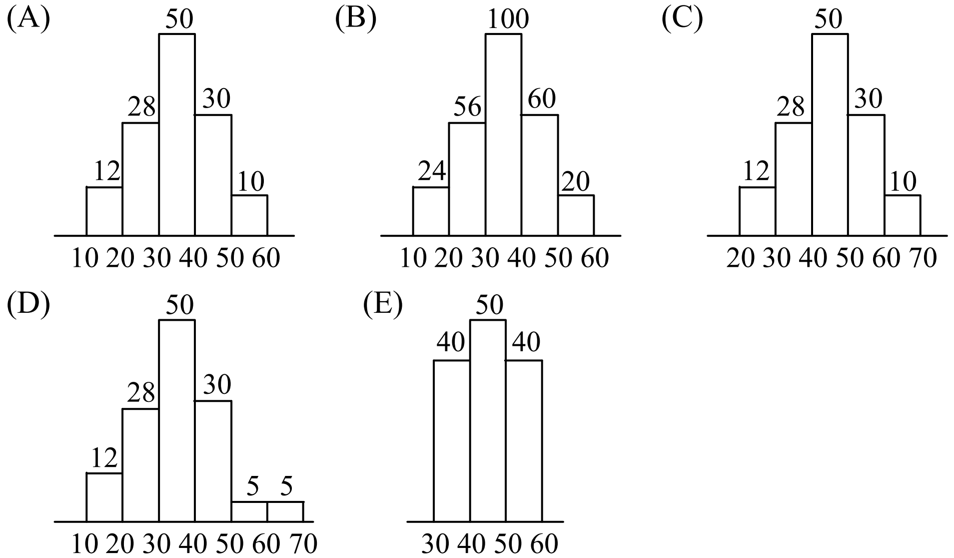
【97 指考乙】

解

標準差 $6.1(\text{英吋})=6.1\times 25.4=154.94(\text{毫米})$

故選(D)

7.__________下列五個直方圖表示的資料，何者之標準差最大？【94 指考乙】



解

標準差不受平移的影響，所以$\sigma_A=\sigma_B=\sigma_C$

又 D 選項圖較分散，可知$\sigma_D>\sigma_A$

E 選項較集中，則$\sigma_D>\sigma_A>\sigma_E$

故選(D)

8.__________考慮下列四組數據：$A:1,2,3,4,5$；$B:2,4,6,8,10$

$C:996,997,998,999,1000$；$D:1,4,9,16,25$；

其標準差分別為σ_A，σ_B，σ_C，σ_D，則下列選項哪些是正確的？

(A)$\sigma_B=2\sigma_A$ (B)$\sigma_C=\sigma_A$ (C)$\sigma_C>\sigma_A$ (D)$\sigma_D>\sigma_A$ (E)$\sigma_D=\sigma_B$

解

(A)○：B 數據是 A 數據的兩倍，所以$\sigma_B=2\sigma_A$

(B)○(C)×：對於所有 C 數據皆符合$c_i-995=a_i$，所以$\sigma_C=\sigma_A$

(D)○：對於所有 D 數據皆符合$d_i=a_i^2$，且 A 數據皆為大於 1 的數，所以$\sigma_D>\sigma_A$

(E)×：D 數據比 B 數據更分散，所以$\sigma_D>\sigma_B$

故選(A)(B)(D)

9.___________測量一物件的長度 9 次，得其長(公尺)為 $2.43, 2.46, 2.41, 2.45, 2.44, 2.48, 2.46, 2.47, 2.45$，將上面的數據每一個都乘以 100，再減去 240 得一組新數據為 $3, 6, 1, 5, 4, 8, 6, 7, 5$，試問下列選項何者為真？

(A)新數據的算數平均數為 5　(B)新數據的標準差為 2

(C)原數據的算數平均數為 2.45　(D)原數據的標準差為 0.2

(E)原數據的中位數為 2.45　【88 學測】

解

設舊數據為 X 數據，新數據為 Y 數據

所以 X，Y 數據的關係為 $y_i = 100x_i - 240$

(A)○：$\mu_y = \dfrac{3+6+1+5+4+8+6+7+5}{9} = 5$

(B)○：$\sigma_y = \sqrt{\dfrac{(-2)^2+1^2+(-4)^2+0^2+(-1)^2+3^2+1^2+2^2+0^2}{9}} = 2$

(C)○：$5 = 100\mu_x - 240 \Rightarrow \mu_x = 2.45$

(D)○：$2 = 100\sigma_x \Rightarrow \sigma_x = 0.02$

(E)○：將新數據由小排到大，得 $1, 3, 4, 5, 5, 6, 6, 7, 8$，可知新數據中位數為 5

則舊數據的中位數為 $\dfrac{5+240}{100} = 2.45$

故選(A)(B)(C)(D)(E)

10.某地區從早上 7:00 到晚上 7:00 的 12 個小時內，每隔一小時測量的雨量(單位 mm)如下：

$6, 8, 9, 20, 25, 32, 20, 18, 10, 6, 5, 1$，則第一四分位數為___________，中位數為___________，第三四分位數為___________。

解

$t = 12 \times \dfrac{25}{100} = 3$，$Q_1 = \dfrac{x_{(3)} + x_{(4)}}{2} = \dfrac{6+6}{2} = 6$

$t = 12 \times \dfrac{50}{100} = 6$，$Me = \dfrac{x_{(6)} + x_{(7)}}{2} = \dfrac{9+10}{2} = 9.5$

$t = 12 \times \dfrac{75}{100} = 9$，$Q_3 = \dfrac{x_{(9)} + x_{(10)}}{2} = \dfrac{20+20}{2} = 20$

[進階題]

11.某次考試中，十位學生成績的平均分數為 56 分，標準差為 4 分，若已知其中八位學生的成績分別為$50,52,53,54,56,57,60,61$，則另兩位學生的分數為________。

解

設兩位同學分數各為x,y

因為平均數$\mu=\dfrac{50+52+53+54+56+57+60+61+x+y}{10}=56$，整理得$x+y=117$

又標準差$\sigma=\sqrt{\dfrac{(-6)^2+(-4)^2+(-3)^2+(-2)^2+0^2+1^2+4^2+5^2+(x-56)^2+(y-56)^2}{10}}=4$

整理得$(x-56)^2+(y-56)^2=53$

將$y=117-x$代入，可得$x=63,54$，則$y=54,63$

所以另外兩位同學分數為$63,54$

12.假設數據x_1,x_2,x_3,x_4,x_5的標準差是$\sqrt{19}$，則函數 $f(x)=(x-x_1)^2+(x-x_2)^2+(x-x_3)^2+(x-x_4)^2+(x-x_5)^2$的最小值為___________。

解

展開$f(x)$可得$f(x)=5x^2-2(x_1+x_2+x_3+x_4+x_5)x+(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)$

整理得$f(x)=5[x-(\dfrac{x_1+x_2+x_3+x_4+x_5}{5})]^2+\triangle$

可知當$x=\dfrac{x_1+x_2+x_3+x_4+x_5}{5}=\mu$，$f(x)$有最小值

又$\sigma=\sqrt{\dfrac{(x_1-\mu)^2+(x_2-\mu)^2+(x_3-\mu)^2+(x_4-\mu)^2+(x_5-\mu)^2}{5}}=\sqrt{19}$

整理得$(x_1-\mu)^2+(x_2-\mu)^2+(x_3-\mu)^2+(x_4-\mu)^2+(x_5-\mu)^2=95$

$f(x)$的最小值為 95

13.________有 10 個數據如下：$1,1,3,3,5,5,6,6,6,6$。今從此 10 個數據中任取 1 數刪除，試問剩餘的 9 個數據中，下列選項何者不變？
(A)算術平均數　(B)中位數　(C)眾數　(D)第一四分位數　(E)標準差

解

(A)×：$\mu=\dfrac{1+1+3+3+5+5+6+6+6+6}{10}=3.6$，所以不論刪去哪個數，平均都會變

(B)○：原數據中位數$Me=\dfrac{x_{(5)}+x_{(6)}}{2}=\dfrac{5+5}{2}=5$，刪去一個數據後$Me'=x_{([4.5]+1)}=x_{(5)}$
不論刪去哪個數都會得$x_{(5)}=5$，所以中位數不變

(C)○：6 出現 4 次，即使刪了一個 6，6 出現次數還是最多，所以眾數不變

(D)○：同中位數，$Q_1=x_{(3)}=Q_1'$，不管刪了哪個數都會得$x_{(3)}=3$

(E)×：不管刪去哪個數都會影響數據的分散程度

故選(B)(C)(D)

14. 某次數學考試，甲班 20 位同學之平均成績為 60 分，標準差為 10 分，乙班 30 位同學之平均成績為 70 分，標準差為 5 分，現將甲、乙兩班的成績合併計算，試求：

(1)合併後的平均成績__________分。

(2)合併後的標準差___________分。

解

設甲班同學分數為 X 數據：$x_1, x_2, x_{3,} \cdots, x_{20}$

乙班同學分數為 Y 數據：$y_1, y_2, y_{3,} \cdots, y_{30}$

(1) $\mu = \dfrac{20\times 60 + 30\times 70}{50} = 66$

(2) $\sigma_x = \sqrt{\dfrac{{x_1}^2 + {x_2}^2 + \cdots + {x_{20}}^2}{20} - {\mu_x}^2} \Rightarrow {x_1}^2 + {x_1}^2 + \cdots + {x_{20}}^2 = 20(100 + 60^2)$

$\sigma_y = \sqrt{\dfrac{{y_1}^2 + {y_2}^2 + \cdots + {y_{30}}^2}{30} - {\mu_y}^2} \Rightarrow {y_1}^2 + {y_1}^2 + \cdots + {y_{20}}^2 = 30(25 + 70^2)$

$\sigma = \sqrt{\dfrac{20(100 + 60^2) + 30(25 + 70^2)}{50} - 66^2} = \sqrt{79} \approx 8.89$