第三十五單元 單變量數據分析

統計提供決策，做決策需要數據。當我們只對某一個變數（ 或稱變量 ）感興趣，例如只想對身高做分析，這種只有一個變數（ 即只有一行 ）數據的分析稱爲**單變量數據分析**（ 或稱爲一維數據分析)，相對的，如果對兩個變數間的關係感興趣，例如研究身高、體重的關係，稱爲**雙變量數據分析**（ 或稱爲二維數據分析)，如果對多個變數間的關係感興趣稱爲**多變量數據分析**（ 或稱爲多維數據分析)。

(甲)統計圖表

當我們遇到一個變數有幾十個甚至上百個數據時，很難由數據中看出整個數據所提供的訊息，必須將數據“簡化”後才能變成有用的“資訊”。最直接有效的方法是以圖表來描述數據，圖表讓我們很快速、很容易看出數據的分布狀況，我們常聽到“一張圖表”勝過“千言萬語”。統計圖、次數分配表及各種統計量都是用來做數據整理簡化的工作。

(1)數據型態的分類：

統計圖表與統計量等數據分析方法的採用，決定於數據的型態。數據型態基本上依其性質分成**離散型**與**連續型**兩類。兩者的差別在於連續型數據兩個數值間可以再插入無限多個數據，而離散型數據則否。

(a) 離散型數據

如性別、血型、宗教信仰、職業等是分類的**計數數據**，稱爲**離散型數據**。又如學生對學校運動設備滿意度以“1”表示非常不滿意、“2”表示不滿意、“3”表示沒意見、“4”表示滿意、“5”表示非常滿意，依序代表五種不同滿意度的 1，2，3，4，5 是離散數據，且它們之間有次序概念，稱爲**次序數據**。又如比賽的名次冠軍、亞軍、季軍、殿軍也是次序的數據。至於一般分類數據，如性別“男生”、“女生”它們之間無大小之分或如宗教信仰“佛教”、 “道教”、“基督教”、“天主教”、“其他”等，也是沒有次序，稱爲**名目數據**。

(b) 連續型數據

如學生的身高、體重、數學成績等的**計量數據**稱爲**連續型數據**。一般如產品之容量、重量、長度等量測所得的數據也都是連續型數據。連續型數據應是小數點以下很多位數，但爲了簡潔，常只取小數點後一、兩位數，甚至到整數部分，如我們常說身高 172 公分、體重 65 公斤等，但它們仍視爲連續型數據。連續型數據經過處理也可轉換成離散型數據，例如：將身高做分組，165 公分以下爲第一組，身高在 165 公分到 175 公分爲第二組，175 公分以上爲第三組。

單變量數據分析常用統計圖表與統計量表達，常用的統計表爲**次數分配表**，常用的統計圖形有**長條圖**、**直方圖**與**圓餅圖**（ 或稱圓面積圖 ）等。

(2)離散型數據圖表

對於離散型數據，將介紹長條圖、圓餅圖兩種圖形。

我們以大華高中高一甲班 50 位學生的血型、滿意度為例，分別畫出長條圖、圓餅圖。

整理 50 位學生的血型次數分配表、滿意度次數分配表如下。

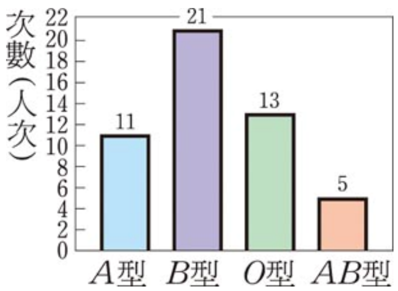
大華高中高一甲班 50 位
學生血型次數分配表

組別	次數	百分比
A 型	11	22
B 型	21	42
O 型	13	26
AB 型	5	10

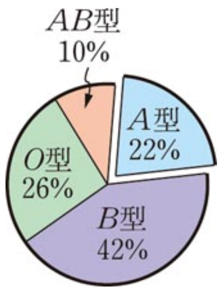
大華高中高一甲班 50 位
學生滿意度次數分配表

組別	次數	百分比
非常不滿意	17	34
不滿意	12	24
沒意見	6	12
滿意	6	12
非常滿意	9	18

因血型是名目數據，橫坐標次序是依類別代號次序由左到右排列，縱坐標為次數，長方形是分開的，圖形如下。



大華高中高一甲班 50 位
學生血型長條圖

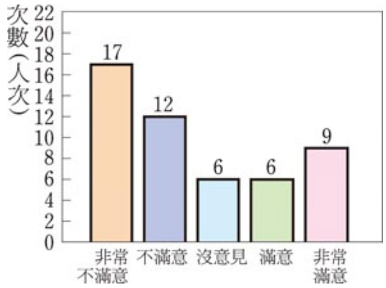


大華高中高一甲班 50 位
學生血型圓餅圖

對於類別數據也可以圓餅圖來表示，每一類畫一扇形，次數愈多的類別，扇形的圓心角愈大，也就是次數與圓心角成正比，上右圖是 50 位學生血型的圓餅圖。

由上面長條圖或圓餅圖看出 50 位學生中血型以 B 型人數最多，佔 42%。

對於次序的數據長條圖橫坐標是依由小到大由左到右依序排列，例如要畫大華高中高一甲班 50 位學生對學校運動設備滿意度長條圖，因滿意度是次序的數據，長條圖橫坐標是對學校運動設備滿意度類別，依非常不滿意到非常滿意，由左到右依序排列，縱坐標為次數，其高度與該類別次數成正比，長方型是分開的，圖形如右。



大華高中高一甲班 50 位學生
對學校運動設備滿意度長條圖

(3)連續型數據圖表

對於連續型的數據，其圖形畫法最常用的是直方圖、次數分配折線圖與累積次數分配折線圖等。

◆ 直方圖

直方圖與長條圖都是由長方形所構成，但有三個差別：

(i) 直方圖通常處理連續型數據，長條圖通常處理離散型數據。

(ii) 直方圖的橫坐標由小而大由左到右依序排列；反之，如爲名目數據的長條圖，次序的安排常依使用者而定。

(iii) 直方圖畫出的長方形是緊連在一起；反之，長條圖畫出的長方形是分開的。

我們以大華高中高一甲班 50 位學生身高爲例，說明直方圖畫法。

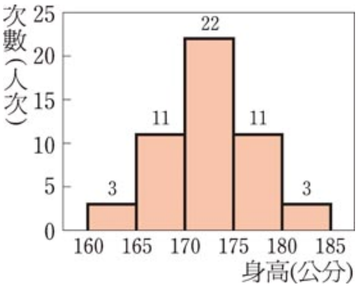
畫直方圖前先要將數據作成次數分配表後，才能依各組次數畫圖。例如身高以 5 公分爲組距，分成 5 組的次數分配表如右。

註：每組含下界不含上界。即 165 公分屬於第二組不屬於第一組。

大華高中高一甲班 50 位
學生身高次數分配表

組別	組界	次數
第一組	160～165	3
第二組	165～170	11
第三組	170～175	22
第四組	175～180	11
第五組	180～185	3

在直方圖中，橫坐標是各組身高的組界，而直方圖中的長方形是連接在一起，縱坐標是次數，也就是長方形高度是該組的次數。

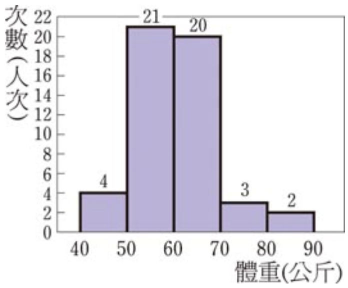


大華高中高一甲班 50 位
學生身高直方圖

右圖有中間高兩邊低、對稱且只有一個高峰（稱爲單峰）的現象，但並不是所有直方圖都如此。例如：大華高中高一甲班 50 位學生體重以 10 公斤爲組距，分成 5 組，其次數分配表與直方圖如下。

大華高中高一甲班 50 位
學生體重次數分配表

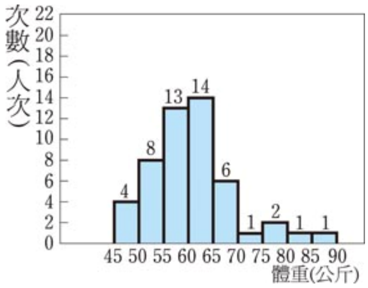
組別	組界	次數
第一組	40～50	4
第二組	50～60	21
第三組	60～70	20
第四組	70～80	3
第五組	80～90	2



大華高中高一甲班 50 位學生
體重直方圖
(以 10 公斤爲組距分成 5 組)

如將體重改以 5 公斤爲組距，分成 9 組，其次數分配表與直方圖如下。

組別	組界	次數
第一組	45～50	4
第二組	50～55	8
第三組	55～60	13
第四組	60～65	14
第五組	65～70	6
第六組	70～75	1
第七組	75～80	2
第八組	80～85	1
第九組	85～90	1



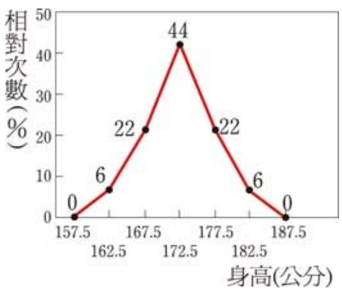
大華高中高一甲班 50 位
學生體重直方圖
（以 5 公斤爲組距分成 9 組）

上面體重直方圖是不對稱的，60 公斤附近人數較多，但由於少數幾位（2 位）學生的體重在 80 公斤以上，造成圖形右邊尾巴拉的很長，我們稱此種直方圖爲偏向右方（簡稱右偏或稱正偏）。值得注意的是分組方式不同直方圖圖案（或長相）也會不同。

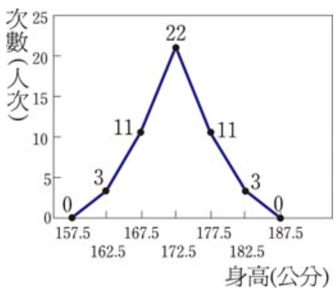
◆ 次數分配(相對次數)折線圖

次數分配折線圖是由直方圖上每個長方形頂邊中點連成的折線所組成，如前述大華高中學生身高直方圖所對應的次數分配折線圖如下圖。

次數分配折線圖中左、右兩端各增加一組，分別爲 155～160 與 185～190，其對應高度都是 0。　若縱坐標不是次數而是相對次數（即$\frac{各組次數}{總數}$）所得的折線圖稱爲相對次數分配折線圖。以大華高中高一甲班 50 位學生身高爲例，其相對次數分配折線圖如圖所示。



大華高中高一甲班 50 位
學生身高次數分配折線圖



大華高中高一甲班 50 位學生
身高相對次數分配折線圖

◆ 累積次數(相對累積)分配折線圖

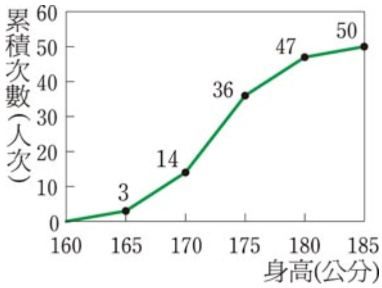
以大華高中高一甲班 50 位學生身高爲例，說明累積次數分配折線圖的畫法，首先是將次數分配表中再增加累積次數及相對累積次數如下。

大華高中高一甲班 50 位學生身高的累積次數與相對累積次數分配表

組別	組界	次數	累積次數	相對累積次數（％）
第一組	160～165	3	3	6
第二組	165～170	11	14	28
第三組	170～175	22	36	72
第四組	175～180	11	47	94
第五組	180～185	3	50	100

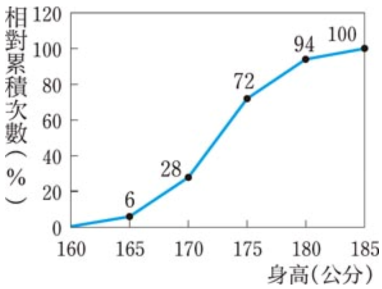
其中相對累積次數是累積次數除以總人數 50 次數，分配表製作好後，可進一步做爲累積次數分配折線圖與相對累積次數分配折線圖的依據。

累積次數分配折線圖是以各組上界爲橫坐標，該組累積次數爲縱坐標的點連成線段而成。例如：以大華高中高一甲班 50 位學生身高爲例，坐標點分別爲 (160，0)，(165，3)，(170，14)，(175，36)，(180，47)，(185，50)，累積次數分配折線圖如下。



大華高中高一甲班 50 位學生身高累積次數分配折線圖

由上圖可得，身高超過 180 公分的只有 3 位，而身高小於或等於 170 公分的有 14 位。相對累積次數分配折線圖的畫法，與累積次數分配折線圖相似，相對累積次數分配折線圖是以各組上界爲橫坐標，該組相對累積次數爲縱坐標，再將這些坐標點連成線段而成，例如：以大華高中高一甲班 50 位學生身高爲例，各坐標點分別爲 (160，0)，(165，6)，(170，28)，(175，72)，(180，94)，(185，100)，其相對累積次數分配折線圖如下。



大華高中高一甲班 50 位學生身高相對累積次數分配折線圖

所以相對累積次數分配折線圖與累積次數分配折線圖形狀相似，只是縱坐標累積次數改爲相對累積次數。

所謂的統計量就是由一組資料所計算出的單一數值，接下來要討論的統計量可以分成兩種形式：

(1)表達資料集中情形的統計量：
用來提供資料的「中心點」、「代表值」、或是出現頻率最多的某個數據等。
(2)表達資料分散程度的統計量：
對於一組資料而言，我們用這種統計量來表達此組資料的分散程度。

(乙)表達資料集中趨勢的統計量

表達資料集中情形的統計量有**算術平均數、加權平均數、中位數、眾數**等。
(1)算術平均數：
(a)定義：

設 n 個實數值分別為 $x_1,x_2,\ldots\ldots x_n$，則其算術平均數$\mu=\frac{1}{n}(x_1+x_2+\cdots+x_n)$。

(b)算術平均的特性：
①在任一數列中，各項數值與其算術平均數之差的總和為零。
②公式簡單，計算容易，適於代數處理。
③易受極端度量影響。
④應用前需考慮：所有資料的數值十分集中、各數值具有同等的重要性。
(c)利用 Excel 計算算術平均數：
我們可以利用 Excel 來計算算術平均數：

= =AVERAGE(D2:D6)

C	D	E
	資料	
	32	
	46	
	34	
	57	
	83	
	平均	
	50.4	

[例題1] (算術平均數的特點)
若 $x_1,x_2,\ldots,x_n$ 之算術平均數為μ_X，若 y_i=Ax_i+B(常以 Y=AX+B 表示)而 $y_1,y_2,y_3,\ldots,y_n$ 之算術平均數為μ_Y，證明：
(1) $\sum_{i=1}^{n} x_i=n\mu_X$ (2) $\sum_{i=1}^{n}(x_i-\bar{x})=0$ (3) $\mu_Y=A\mu_X+B$。

(練習1) 某班段考的數學成績經統計後，得到平均分數爲 48 分，而且最高分也只有 60 分，因此老師決定將每人的成績乘以 1.5 後，再加 10 分，求經此調整後，平均分數爲__________分。Ans：82

(練習2) 某一組資料爲 N_1 個，其平均值 $\bar{y}_1$；另一組資料 N_2 個，其平均值 $\bar{y}_2$，則此兩組混合後之平均值爲(A) $\dfrac{\overline{y_1}+\overline{y_2}}{2}$ (B) $\dfrac{\overline{y_1}+\overline{y_2}}{N_1+N_2}$ (C) $\dfrac{N_1\overline{y_1}+N_2\overline{y_2}}{N_1+N_2}$ (D) $\dfrac{N_1\overline{y_1}+N_2\overline{y_2}}{2}$。Ans：(C)

(2)加權平均數：

設實數 x_i 的權數爲 w_i，i=1,2,3,….,n，則 $x_1,x_2,\ldots\ldots x_n$ 的加權平均數 W 定義爲

$W=\sum_{i=1}^{n} x_i w_i$ 。

(a)爲什麼要使用加權平均數？

統計資料中，如果每一筆資料的重要性不同，則必須使用加權來處理平均數。

例如：學生各科成績的重要性依上課時數的不同而異，此時若以各科成績的算術平均數作爲學生的學期總平均成績，顯然不適當，於是爲了正確的表達評量標準，必須考慮各科的授課時數而採用加權的方式來處理。

(b)加權平均數的適用時機爲何？

若各科平均數資料的**重要性**不同，則必須對各筆統計資料分別賦予一個對應的權數，再利用加權平均數來表示母群體的性質。

(3)中位數(Median)

一組資料的中位數 Me 代表大於 Me 或小於 Me 的資料數目，均各佔 50%。

(a)未分組資料：設資料有 n 個數據

n 爲奇數時，中位數是指按大小順序排列後之第 $\dfrac{n+1}{2}$ 個數。

n 爲偶數時，中位數是指按大小順序排列後之第 $\dfrac{n}{2}$個數與第 $\dfrac{n}{2}+1$ 個數的平均數。

(b) 由已分組資料求中位數

Step1：將 n 個數值先作出「次數分配表」，再求出「以下累積次數表」

組別	次數 f	以下累積次數 C
L_1~U_1	f_1	$C_1=f_1$
………	……….	………
L_{i-1}~U_{i-1}	f_{i-1}	$C_{i-1}=f_1+f_2+f_3+\ldots+f_{i-1}<\frac{n}{2}$
L_i~U_i	f_i	$C_i=f_1+f_2+\ldots+f_{i-1}+f_i\geq\frac{n}{2}$
……..	……..	
L_k~U_k	f_k	
…總計…	n	

Step2：找出 Me 所在的組別

Step3：利用比例關係求 Me⇒ $\dfrac{Me-L_i}{U_i-L_i}=\dfrac{\frac{n}{2}-C_{i-1}}{C_i-C_{i-1}}$ 。

(c)中位數的適用時機，說明如下：

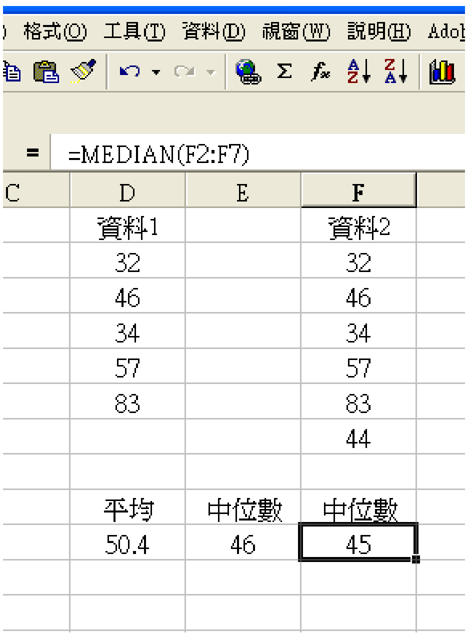
①若某一筆資料比中位數大，則可知道該筆資料位在母群體的上半部內，
若某一筆資料比中位數小，則可知道該筆資料位在母群體的下半部內。

②中位數是整個母群體的中心，不受資料中有特別大的數或特別小的數的影響。

③適合常態分配的母群體。

(d)利用 Excel 計算中位數：

我們可以利用 Excel 計算中位數：



[例題2] 設甲、乙兩組第一次月考數學分數均不同，如下：

甲組：31，40，45，46，49，50，53，55，57，58，60，70，75

乙組：50，53，60，67，68，69，72，76，80，83

(1)甲組之中位數為__________ (2)乙組之中位數為__________。
Ans：(1)53 (2)68.5

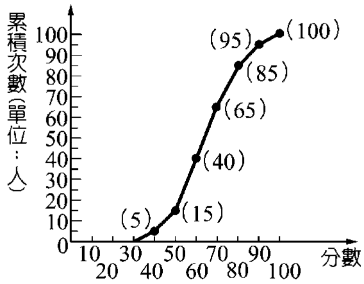
[例題3] 某校三年級學生 120 人參加模擬考試，其成績分布如下：

成績	10～20	20～30	30～40	40～50	50～60
人數	1	5	10	15	23

成績	60～70	70～80	80～90	90～100
人數	30	19	14	3

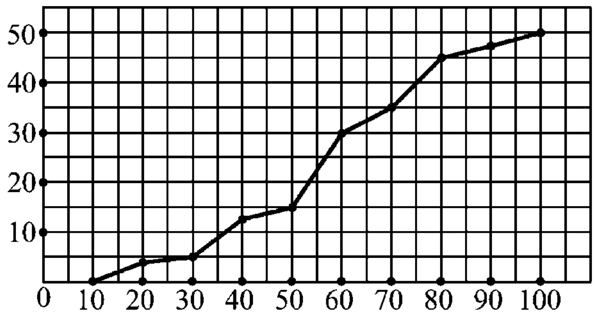
求中位數。Ans：62

(練習3) 中山國小六年級學生 100 人，某次數學考試成績之累積次數分配曲線圖如下（括弧內數字表示累積次數）。



假設各組內之次數都平均分布在組距內，
則：(1)算術平均數為__________；
(2)中位數為__________。
（四捨五入成整數）
Ans：(1)64.5(2)64

(練習4) 某班 50 位同學數學科成績的以下累積次數分配曲線如圖所示，則成績的中位數為__________（取到整數，小數點以下四捨五入）Ans：57



(練習5) 12 位同學在罰球線上投籃 10 次，投中次數分別爲 3,2,3,7,5,3,6,4,1,3,6,8
(1)試求這 12 位同學投中次數的中位數。
(2)若小安投中 6 次，請問這 13 位同學投中次數的中位數。
Ans：(1)3.5 次 (2)4 次

(4)幾何平均數：
對於一組數據，表達數據代表值的統計量除了上述的算術平均數、中位數、眾數外，有時使用幾何平均數。
(a)定義：
一組正數資料 $x_1,x_2,x_3,\ldots\ldots,x_n$ 的幾何平均數(簡寫成 G.M.)，定義爲 G.M.$=\sqrt[n]{x_1x_2\cdots x_n}$ 。

(b)實例：2,4,8,16,32 的幾何平均數爲 $\sqrt[5]{2^{15}}=8$，算術平均數爲$\frac{62}{5}$。

(c)應用：
人類的經濟活動，常常要表達薪資、經濟指標、營業額、投資報酬率等每年的變化率，我們在表達在一段時間內的變化率或成長率的平均值，常以幾何平均數來表示。
例如：某公司去年的銷售量比前年成長 20%，而今年的銷售額比去年衰退 20%，請問這兩年間的平均銷售量爲多少？
[解法一]：用算術平均數，平均銷售量= $\frac{20\%-20\%}{2}=0$。
[解法二]：用幾何平均數
設前年的銷售量爲 A，則去年的銷售量爲 A(1+20%)，
而今年的銷售量爲 A(1+20%)(1−20%)=A(1−4%)=A·96%

換句話說，這兩年來的銷售是呈現衰退現象，如用算術平均數來表示的話，會得出今年的銷售量與前年相同，這與實際的情形並不相同，因此用算術平均數無法表現出實際的成長情形。**如何計算這兩年的平均成長率呢**？

設這兩年的平均成長率爲 x，則去年的銷售量爲 $A(1+x)$，今年的銷售量爲 $A(1+x)^2$
$\Rightarrow A(1+x)^2= A(1+20\%)(1-20\%)$
$\Rightarrow(1+x)^2=1.2\times0.8 \Rightarrow 1+x=\sqrt{1.2\times0.8} \Rightarrow x=\sqrt{0.96}-1=-0.0202$
所以這兩年的平均成長率爲−2.02%
(即銷售量成負成長，平均每年減少銷售量的 2.02%)。
結論：
若 n 年的成長率分別爲 y_1、y_2、y_3、…、y_n，
則這 n 年的平均成長率爲 $\sqrt[n]{(1+y_1)(1+y_2)\cdots(1+y_n)}-1$。

(d)性質：
①幾何平均數計算不易，常對資料取對數後，再求算術平均數。

$\log(\text{G.M.})=\frac{1}{n}(\log x_1+\log x_2+\ldots+\log x_n)$

②算術平均數≥幾何平均數。

③對於有極端值得一組資料，幾何平均數比算術平均數有穩定性(即較不敏感)。

[例題4] 台灣地區民國 80 年到 87 年這 8 年的經濟成長率分別爲 7.55，6.76，6.32，6.54，6.03，5.67，6.77，4.65(單位：%)，請問台灣地區這 8 年平均經濟成長率爲多少？ Ans：6.2832%

[解法]：

設這 8 年平均經濟成長率爲 x

$(1+x)^8=1.0755\times1.0676\times1.0632\times1.0654\times1.0603\times1.0567\times1.0677\times1.0465$

$\Rightarrow x=(1.0755\times1.0676\times1.0632\times1.0654\times1.0603\times1.0567\times1.0677\times1.0465)^{\frac{1}{8}}-1$

$\Rightarrow x=0.062832=6.2832\%$

[例題5] 某公司民國 85 年營業額爲 4 億元，民國 86 年營業額爲 6 億元，該年的成長率爲 50%。87、88、89 年三年的成長率皆相同，且民國 89 年的營業額爲 48 億元。則該公司 89 年的成長率爲______%。

Ans：100 (91 學科)

(練習6) <u>大明</u>開設一公司，連續三年的成長率依序爲−10%，20%，60%，則此公司三年的年成長率平均值爲__________。Ans：20%

(練習7) 新成立的某家公司，四年來的營業額依次爲 242，271，264，306（單位：百萬元），則每年平均成長率爲________。(已知 $\log 1.21 = 0.0828$，$\log 1.53 = 0.1847$，$\log 1.082 = 0.0340$)Ans：8.2%

(丙)表達資料分散趨勢的統計量

表達資料分散程度的統計量有**全距、四分位距、標準差**等。

(1)全距：

(a)一群統計資料中，最大數據與最小數據之差，叫做全距，以 R 表之。

(b)求法：

未分組資料⇒將資料由小而大排列，$x_1\le x_2\le\ldots\ldots\le x_n$，則 R=最大−最小。

已分組資料⇒R=(最大一組的上限) − (最小一組的下限)

(c)特性：

優點：容易了解，計算簡單。

缺點：當數值資料較集中時，利用全距可以清楚表達離散的程度，但中間數值的變動情況不易得知，所以資料不集中時不適用。

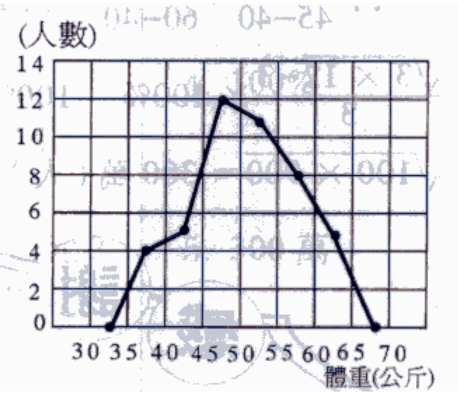
[例題6] 右圖是某班同學的次數分配折線圖，求其全距。

Ans：30

[解法]：

因爲 30~35 與 65~70 的人數均爲 0，

故 R=65–35=30。



(2)變異數與標準差

(a)想法：

①差異量數中的全距與四分位數，都是根據統計資料中某特定兩點的距離而得，無法代表全部資料的分散情形，且全距易受抽樣的影響，四分位差忽視百分之五十的資料，故都不是很好的差異量數，因此較少被使用。

②標準差是以全部資料的算術平均數爲中心，計算全部資料與算術平均數的平均距離，代表整個資料的分散情形。

(b)平均離差：

要表達一組資料的分散程度，我們很自然想到以資料離中心點多遠來表示。

設有一組資料 $x_1,x_2,x_3,\ldots\ldots,x_n$，平均數爲 $\bar{x}$，

則第 i 筆資料 x_i 的離均差定義爲 $\boxed{x_i-\bar{x}}$。

且自然容易想說只要將全部的離均差相加再除以 n，可了解資料的分散程度。

要就是說只要算出 $\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})}{n}$ 即可知道**平均離差**爲何。但不幸的是，

因爲 $n\bar{x}=\sum_{i=1}^{n}x_i$ ，也就是說 $\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})}{n}$ =0。因此平均離差無法提供資料分散的程度。

(c)平均絕對離差：

將離差改成絕對離差$|x_i-x|$，再算出絕對離差的平均數，就可以看出這筆資料的平均分散程度。

定義：一組資料 $x_1,x_2,x_3,\ldots\ldots,x_n$ 的**平均絕對離差**為 $\dfrac{\sum_{i=1}^{n}|x_i-\bar{x}|}{n}$ 。

例如：

甲地區的早、中、晚的氣溫為 10,20,30；乙地區為 19,20,21 平均溫度是一樣的。

甲的平均絕對離差為$\frac{1}{3}(|-10|+|10|+|0|)=\frac{20}{3}$，乙的平均絕對離差為$\frac{1}{3}(|-1|+|1|+|0|)=\frac{2}{3}$。

⇒甲地區比乙地區的溫差變化大。

平均絕對離差是一個很自然就會想到當作測量一組資料有多分散的指標，但可惜的是**絕對值**在代數的運算上非常的麻煩，於是將絕對值變成平方，即為變異數的概念。

(3)變異數與標準差

一組資料 $x_1,x_2,x_3,\ldots\ldots,x_n$ 的**變異數**σ^2 為 $\dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$ 。μ為資料的算術平均數

由於變異數的單位是資料單位的平方，(如本來單位是 cm，但變異數為 cm^2)，它必須開方後才能與平均數、中位數等「集中」統計量做加、減等運算，因此常以變異數開方來表示資料的分散程度，稱之為**標準差**。

標準差$\sigma=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2}=\sqrt{(\frac{1}{n}\sum_{i=1}^{n}x_i^2)-\mu^2}$

(4)標準差的特性：

(a)以算術平均數為中心的標準差，較任何其它的平均數為中心的標準差小。

(b)標準差的特性與算術平均數相同。

(c)標準差易於做代數運算。

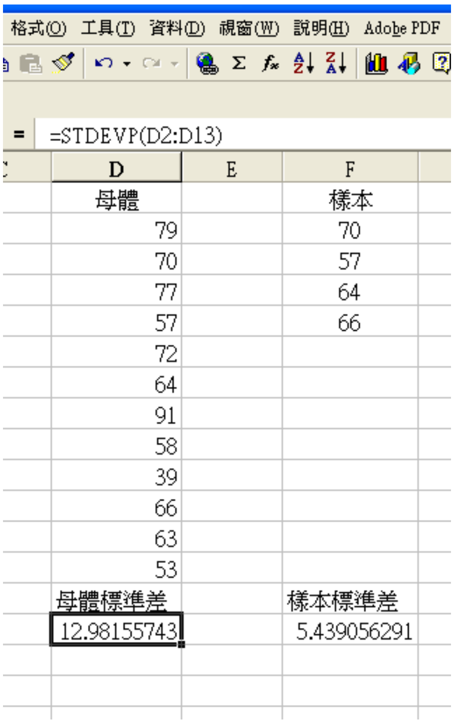
(5)資料轉換時統計量的關係：

設原資料以 **X** 表示，將 **X** 中的每筆資料乘以 ***a*** 再加上 ***b***，形成新的資料 **Y**，我們將其寫成 **Y=*a*X+*b***。

則 **Y** 的平均量數(算術平均數、中位數、加權平均數)= ***a*(X** 的平均量數**)+*b***

Y 的差異量數(全距、四分位距、標準差)=$|a|$**(X** 的差異量數**)**

(6)利用 Excel 計算標準差：



D	E	F
母體		樣本
79		70
70		57
77		64
57		66
72		
64		
91		
58		
39		
66		
63		
53		
母體標準差		樣本標準差
12.98155743		5.439056291

D	E	F
母體		樣本
79		70
70		57
77		64
57		66
72		
64		
91		
58		
39		
66		
63		
53		
母體標準差		樣本標準差
12.98155743		5.439056291

[例題7] 由台北市選民抽出 n 位，訪問他們是否要投票給候選人甲，結果有 k 位選民要投票給候選人甲，設第 i 位受訪者資料爲 x_i，即若第 i 位受訪者投給候選人甲，則 $x_i=1$，否則 $x_i=0$，求候選人甲樣本得票率與標準差。

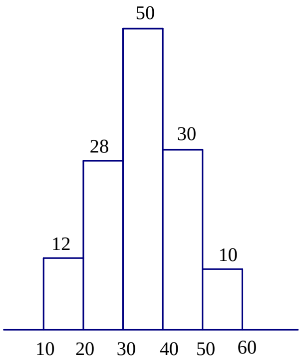
Ans：樣本得票率=$\frac{k}{n}$，標準差=$\sqrt{p(1-p)}$ $(p=\frac{k}{n})$

[例題8] 有 10 個資料 x_1，x_2，……，x_{10}，已知 $\sum_{i=1}^{10} x_i = 155$，$\sum_{i=1}^{10} x_i^2 = 2551$，求其算術平均數爲____________，標準差爲__________。

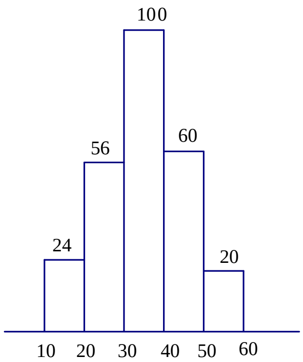
Ans：15.5；3.9

[例題9] 下列五個直方圖表示的資料，何者之標準差最大？ (2005 指定乙)

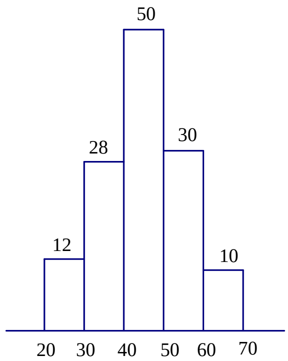
(1)



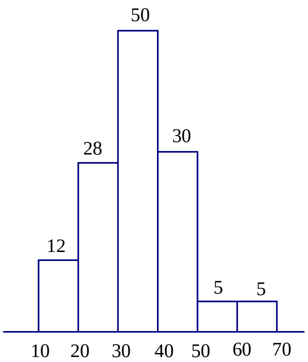
(2)



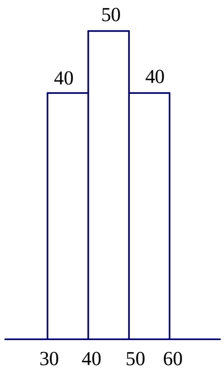
(3)



(4)



(5)



Ans：(4)

[例題10] n 筆數據 x_1，x_2，⋯，x_n 的平均數為 μ_x，標準差為 σ_x，將每筆數據乘 a，再加 b 成另一組數據 y_1，y_2，⋯，y_n，即 $y_i=ax_i+b$，$i=1$，2，⋯，n，令 μ_y，σ_y 為 y_1，y_2，⋯，y_n 的平均數與標準差。

試證：(1) $\mu_y=a\mu_x+b$。

(2) $\sigma_y=|a|\cdot\sigma_x$。

證明：(1) $\mu_y=\dfrac{\sum\limits_{i=1}^{n}y_i}{n}=\dfrac{\sum\limits_{i=1}^{n}(ax_i+b)}{n}=\dfrac{1}{n}(\sum\limits_{i=1}^{n}ax_i+\sum\limits_{i=1}^{n}b)$

$=\dfrac{1}{n}(a\cdot\sum\limits_{i=1}^{n}x_i+nb)=a\mu_x+b$。

(2) $\sum\limits_{i=1}^{n}(y_i-\mu_y)^2=\sum\limits_{i=1}^{n}[(ax_i+b)-(a\mu_x+b)]^2$

$=\sum\limits_{i=1}^{n}a^2(x_i-\mu_x)^2=a_2\cdot\sum\limits_{i=1}^{n}(x_i-\mu_x)^2$，

所以 $\sigma_y=\sqrt{\dfrac{\sum\limits_{i=1}^{n}(y_i-\mu_y)^2}{n}}=\sqrt{\dfrac{a^2\sum\limits_{i=1}^{n}(x_i-\mu_x)^2}{n}}=|a|\sigma_x$。

[例題11] 測量一物件 9 次，得其長度(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 得一組新的數距為 3，6，1，5，4，8，6，7，5，問下列選項何者為真？ (A)新數據的算術平均數為 5 (B)新數據的標準差為 2 (C)原數據的算術平均數為 2.45 (D)原數據的標準差為 0.2 (E)原數據的中位數為 2.45

Ans：(A)(B)(C)(E) (88 學科)

[例題12] 若 $f(x)=\dfrac{1}{n}\sum\limits_{k=1}^{n}(x_k-x)^2$ 為 x 的二次多項式，證明：

當 $x=\mu$ 時，$f(x)$ 有最小值 σ^2。

其中 μ 與 σ 分別為 x_1、x_2、⋯、x_n 的算術平均數與標準差。

[例題13] (合併資料的算數平均數與標準差)

已知 A、B 兩班的人數分別爲 20、30 人，某次考試成績的算術平均數分別爲 75 分、60 分，標準差分別爲 5 分、6 分，求利用母體標準差的公式來計算合併兩班之後的算術平均數與標準差。

【解一】：$\overline{X}=\dfrac{20\times75+30\times60}{20+30}=\dfrac{2\times75+3\times60}{2+3}=66$

$\sigma^2=\dfrac{20[5^2+(75-66)^2]+30[6^2+(60-66)^2]}{20+30}$

$\Rightarrow\sigma^2=\dfrac{20(25+81)+30(36+36)}{50}=\dfrac{2\times106+3\times72}{5}=85.6$

$\therefore\sigma=\sqrt{85.6}\doteqdot9.3$

【解二】：1^o 設 A 班之 20 人之分數分別爲 $x_1,x_2,\cdots\cdots,x_{20}$

則 $\overline{X}=75$，$S_1=5.0$

$\Rightarrow\sum_{i=1}^{20}x_i^2=n_1(\overline{X}^2+\sigma_1^2)=20[75^2+(5.0)^2]=113000$

同理 B 班之 30 人分數如爲 $y_1,y_2,\cdots\cdots,y_{30}$

則 $\overline{Y}=60$，$\sigma_2=6.0$

$\Rightarrow\sum_{i=1}^{30}y_i^2=n_2(\overline{Y}^2+\sigma_2^2)=30[60^2+(6.0)^2]=109080$

$\therefore\sum_{i=1}^{20}x_i^2+\sum_{i=1}^{30}y_i^2=113000+109080=222080$

2^o 平均值 $M=\dfrac{75\times20+60\times30}{20+30}=66$

標準差 $\sigma=\sqrt{\dfrac{222080}{50}-66^2}=\sqrt{85.6}\doteqdot9.25\doteqdot9.3$

【註】：本題之標準差如果改爲樣本標準差 $S=\sqrt{\dfrac{\Sigma(x_i-\overline{X})^2}{n-1}}$

則 $S^2=\dfrac{19\cdot5^2+20(75-66)^2+29\cdot6^2+30(60-66)^2}{20+30-1}$

$=\dfrac{475+1620+1044+1080}{49}=\dfrac{4219}{49}\doteqdot86.1$

$\therefore S=\sqrt{86.1}=9.28\doteqdot9.3$

(練習8) 求下列各數的標準差：2,3,7,8,10　　Ans：$\sqrt{9.2}$

(練習9) 求 1700,1800,1900,2000,2100,2200 之標準差。　Ans：$\dfrac{50\sqrt{105}}{3}$

(練習10) 小安第一次月考六科的平均成績爲 80 分，若已知其中五科的成績爲 68,80,80,80,86，則其成績的標準差爲________分。　Ans：6 分

(練習11) 變量 X 之算術平均數 $\overline{X}=16$，標準差 $S_x=4$，若 $X=4Y-3$，求 Y 之算術平均數 $\overline{Y}$ 與標準差 S_y。Ans：$\dfrac{19}{4}$；1

(練習12)已知 2011 筆數據 $x_1,x_2,\cdots,x_{2011}$ 的算術平均數爲 4，標準差爲 3，則：

(1) $-3x_1+5,-3x_2+5,\cdots,-3x_{2011}+5$的標準差爲何？

(2) $3{x_1}^2+25,3{x_2}^2+25,\cdots,3{x_{2011}}^2+25$的算術平均數爲何？

Ans：(1)9 (2) 100

(練習13) 有一群資料 X：1，1，2，3，5，5，7，8，9，9，另一群資料 Y：2001，2001，2002，2003，2005，2005，2007，2008，2009，2009，則下列何者正確？

(A) X 之中位數＝5　(B) Y 之算術平均數＝2005

(C) Y 之中位數＝2005　(D) X 之標準差 S_x＝3　(E) Y 之標準差 $S_y=S_x$。

Ans：(A)(B)(C)(D)(E)

(丁)資料的標準化

一組數據 x_i，i＝1，2，⋯，n，乘一常數 a 及加一常數 b 而成另一組數據 y_i ($y_i=ax_i+b$)，此種變換稱爲線性變換，線性變換即爲對數據做伸縮(係數 a)與平移(係數 b)。

例如： 5 位學生的成績如下：

35，45，50，55，65，算出平均分數 μ_x 爲 50，標準差 σ_x 爲 10，

如果老師將 5 位學生的成績調高，將每人成績乘 1.2 再加 3 分，

即 $y_i=1.2x_i+3$ ($a=1.2$，$b=3$)，則 5 位成績調整爲 45，57，63，69，81，這就是線性變換。

經線性變換後原數據與新數據的平均數與標準差有下面的關係式：

n 筆數據 x_1，x_2，⋯，x_n 的平均數爲 μ_x，標準差爲 σ_x，將每筆數據乘 a，再加 b 成另一組數據 y_1，y_2，⋯，y_n，即　$y_i=ax_i+b$，i＝1，2，⋯，n，
令 μ_y，σ_y 爲 y_1，y_2，⋯，y_n 的平均數與標準差。則
(1) $\mu_y=a\mu_x+b$。(2) $\sigma_y=|a|\cdot\sigma_x$。

[例題14] 有一組數據 x_1，x_2，⋯，x_{10} 的平均數爲 15，標準差爲 0.3，現在將每筆數據減 15，再除以 0.3，成爲一組新數據 y_1，y_2，⋯，y_{10}，即 $y_i=\dfrac{x_i-15}{0.3}$，i＝1，2，⋯，10。試求 y_1，y_2，⋯，y_{10} 此筆數據的平均數與標準差。

Ans：0，1

最常用的線性變換是將數據變換成平均數為 0，標準差為 1 的新數據，我們稱此種線性變換為**數據標準化**，數據標準化的做法如下：

令 x_i 表第 i 筆原數據，而 x_i' 表第 i 筆新數據，則 $x_i'=\dfrac{x_i-\mu}{\sigma}$ ，

其中 μ ，σ 是原數據的平均數與標準差，由於數據 x_i 的平均數 μ 與標準差 σ 是同單位，所以標準化後的 x_i' 為無單位。

舉例說明數據標準化如下：

例如：5 位學生的成績如下：

35，45，50，55，65，算出平均分數 $\mu=50$，標準差 $\sigma=10$，

第 1 位學生標準化成績為 $x_1'=\dfrac{35-50}{10}=-1.5$，

第 2 位學生標準化成績為 $x_2'=\dfrac{45-50}{10}=-0.5$ ，

其他三位學生標準化成績分別為 0，0.5，1.5。

舉例說明數據標準化的應用：

設某位學生第一次段考數學成績是 57 分，國文成績是 78 分，是否就能說此生的國文考的比數學好呢？事實上，不能單以表面的 78 分比 57 分高，就說國文考的比數學好！如果全班第一次段考數學平均分數是 43 分，標準差是 7 分，而國文平均分數是 73 分，標準差是 5 分，則此生

數學標準化成績是 $\dfrac{57-43}{7}=2$，而此生的國文標準化成績是 $\dfrac{78-73}{5}=1$。

此生數學標準化成績為 2，比國文標準化成績為 1 來的高，表示與其他同學比較，此生數學成績比國文成績好。

(丙)整體描述

要描述一組數值資料，通常會先畫圖，因為圖形可以對數據分布提供最清楚的整體狀況，如下圖所示，當數值資料分組夠多時，將直方圖中各長方形頂邊的中點用平滑曲線相連，就形成次數分配曲線圖。曲線底下的面積就如同直方圖中的長條面積一樣，代表觀測值的次數或是百分比。

在曲線圖中，有幾個我們可以用來了解資料整體分布的統計量：

(1)眾數是「**尖峰點**」，即曲線的最高點所在的數值。

(2)中位數是「**等面積點**」，曲線底下的面積恰好被通過中位數的鉛直線平分。

(3)算術平均數是「**平衡點**」，所有平均數左邊的數值到平均數的距離總和會等於平均數右邊的數值到平均數的距離總和，因此平均數是各數值的重心。

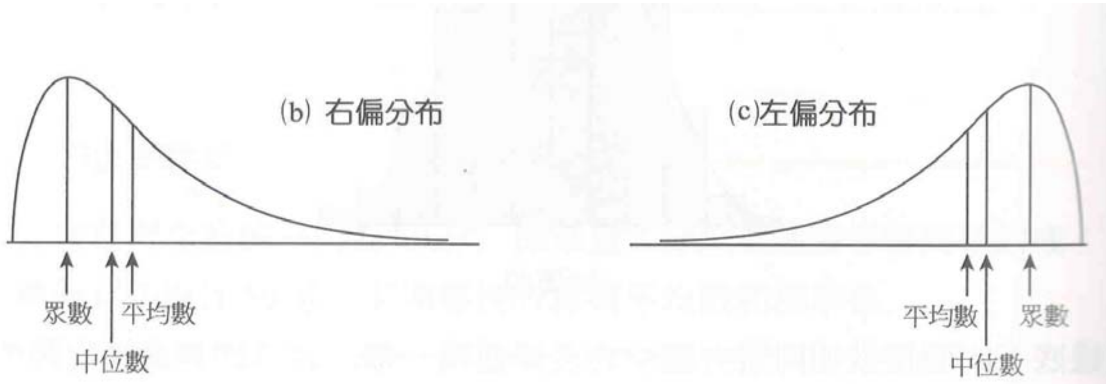
想像數據是疊在翹翹版上的法碼，平均數就是翹翹版的平衡點。

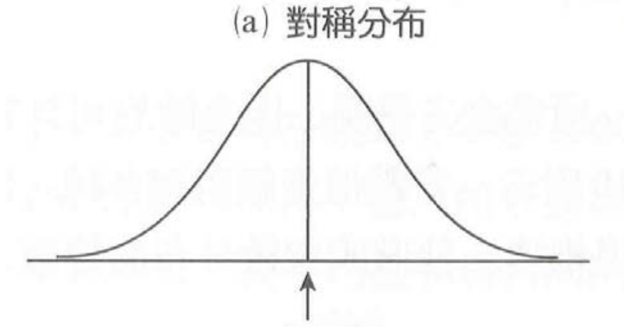
若次數分布曲線的左半與右半形狀大致相等，我們稱該分布**對稱**；若曲線的右邊(數值較大的那一半)延伸出去比左邊遠，則這個分布是**右偏**；若曲線的左邊(數值較小的那一半)延伸出去比右邊遠，則這個分布是**左偏**。

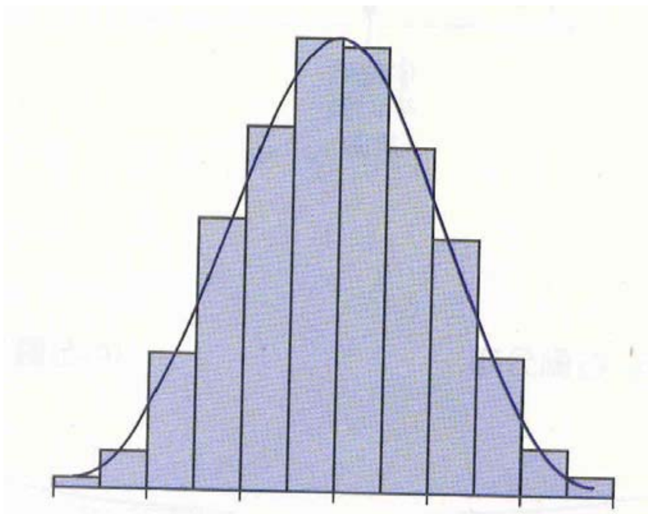
如下圖，一個對稱分布的曲線圖，平均數和中位數是很接近，但是對於右偏或左偏的分布來說，平均數就會被拉往長尾巴的方向(即右偏的平均數會大於中位數，因為平均數易受很大數值影響而變大)。

圖畫好了，應該如何決定用哪一個統計量來描述中心和離度呢？因為平均數和標準差易受極端值影響，而中位數和四分位距教不受影響，一般來說，大致對稱的分布常用平均數和標準差來描述中心和離度，但是要描述偏斜分布時，通常會合併使用最小值、第一四分位數、中位數、第三四分位數、最大值等五個數來描述。

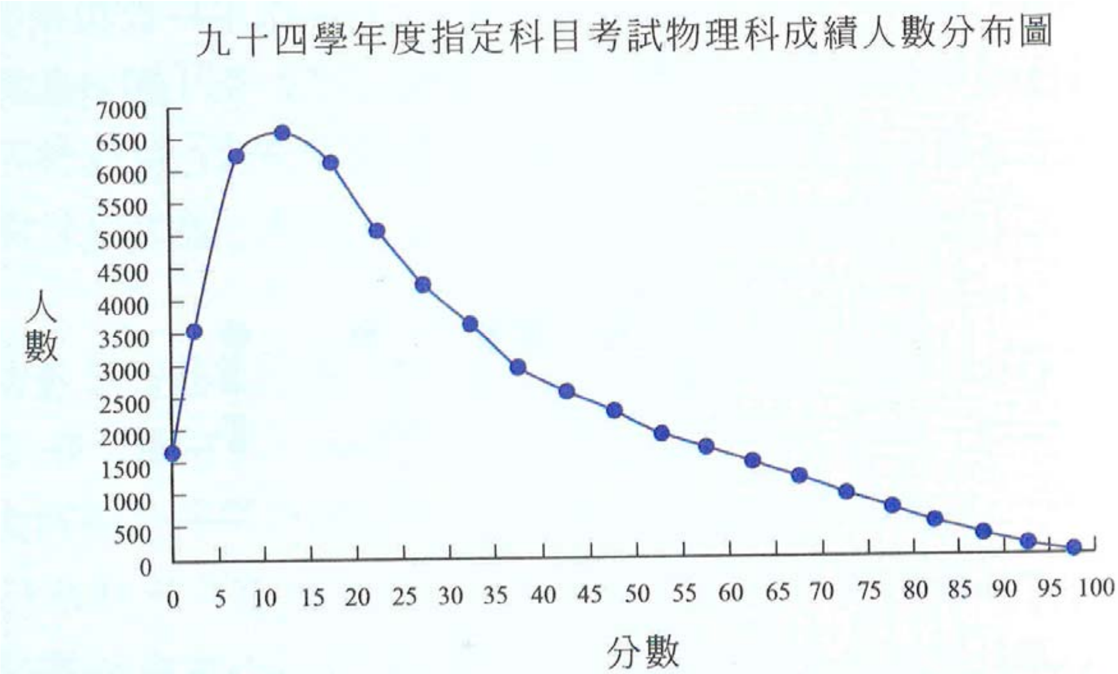
某些特殊的現象會使圖呈現雙峰，例如中小學英文的學習成果，常因城鄉差距而呈現雙峰分布，此時若僅用數字描述並不能說明此現象，圖表或許是最佳的呈現方式。







[例題15] 下圖是 94 年指定考科物理科成績人數分布圖，



(1)此分布型態爲對稱、右偏、左偏？

(2)眾數出現在那一組？

(3)已知該科全部考生成績的均標(第 50 百分位數)爲 23 分，那麼平均分數會大於、等於或小於 23 分？

(練習14) 下圖是某高中學生體重的分布曲線，其中的 3 條鉛直線爲眾數、中位數與平均數三者的位置，下列選項哪些是正確的？

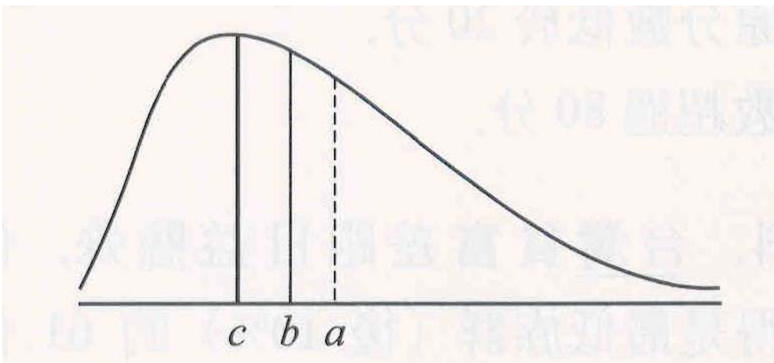
(1)體重的分布爲左偏分布。

(2)a 爲中位數

(3)b 爲平均數

(4)c 爲眾數

Ans：(4)



綜合練習

(1) 調查某新興工業都市的市民對市長施政的滿意情況，依據隨機抽樣，共抽樣男性 600 人、女性 400 人，由甲、乙兩組人分別調查男性與女性市民。調查結果男性中有 36%滿意市長的施政，女性市民中有 46%滿意市長的施政，則滿意市長施政的樣本佔全體樣本的百分比爲_______%。(90 學科)

(2) 在某項才藝競賽中，爲了避免評審個人主觀影響參賽者成績太大，主辦單位規定：先將 15 位評審給同一位參賽者的成績求得算術平均數，再將與平均數相差超過 15 分的評審成績剔除後重新計算平均值做爲此參賽者的比賽成績。現在有一位參賽者所獲 15 位評審的平均成績爲 76 分,其中有三位評審給的成績 92、45、55 應剔除，則這個參賽者的比賽成績爲________分。(96 學科)

(3) 某校想要了解全校同學是否知道中央政府五院院長姓名，出了一份考卷。該卷共有五個單選題，滿分 100 分，每題答對得 20 分，答錯得零分，不倒扣。閱卷完畢後，校方公佈每題的答對率如下：

題號	一	二	三	四	五
答對率	**80%**	**70%**	**60%**	**50%**	**40%**

請問此次測驗全體受測同學的平均分數是

(A)70 分 (B)65 分 (C)60 分 (D)55 分。 (91 指定甲)

(4) 下列 5 組資料(每組各有 10 筆)

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

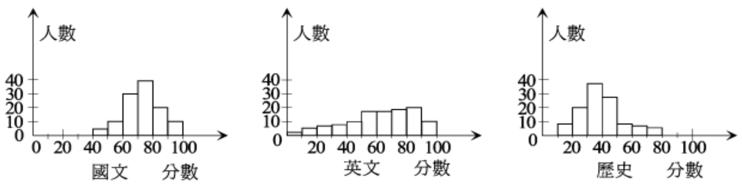
D：1,1,2,2,3,3,4,4,5,5

E：1,2,3,4,5,6,7,8,9,10

試問哪一組資料的標準差最大？

(1)A (2)B (3)C (4)D (5)E (89 學科)

(5)



上列三圖爲某年級國文，英文，歷史三科成績分佈情形的直方圖，根據該圖，下列那些推論是合理的？

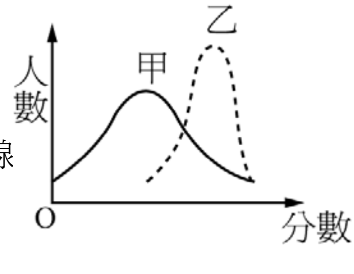
(A)歷史的平均分數比國文的平均分數低 (B)歷史的平均分數最低 (C)英文的眾數最高 (D)英文的標準差比國文的標準差小

(E)英文的標準差最大。

(6) 某年聯考甲、乙兩科成績的直方圖如圖所示，

(由於考生人數眾多，成績分布的直方圖可視爲平滑的曲線

則下列哪些敘述是正確的？



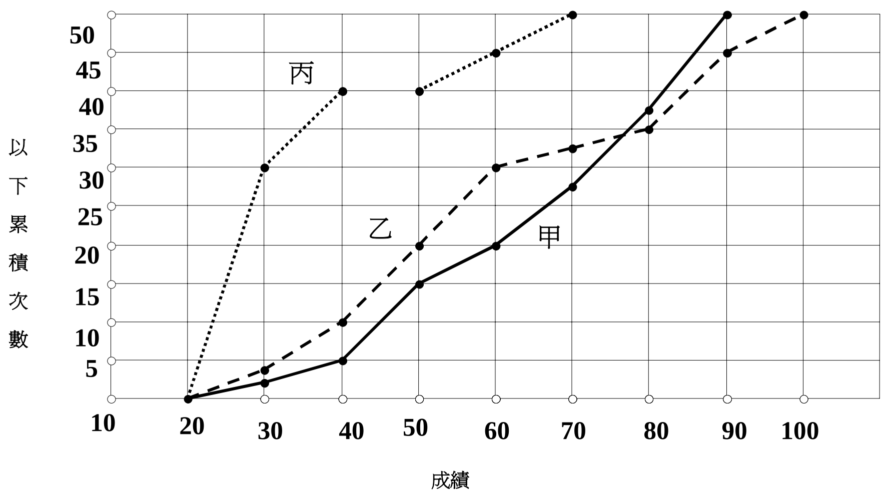
(A)甲的算術平均數比乙的算術平均數大
(B)甲的中位數比乙的中位數大
(C)甲的全距比乙的全距大
(D)甲的標準差比乙的標準差大

(7) 某班有 48 名學生，某次數學考試之成績，經計算得算術平均數 70 分，標準差爲 S 分。後來發現成績登記有誤，某甲得 80 分確誤記爲 50 分，某乙得 70 分確誤記爲 100 分，更正後重算得標準差爲 S_1 分，試問 S_1 與 S 之間，有下列那一種大小關係？

(n 個數值 $x_1 , x_2 , \ldots, x_n$ 的標準差爲 $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2-\overline{x}^2}$ ，而 $\overline{x}=\frac{1}{n}\cdot\sum_{i=1}^{n}x_i$)

(A)$S_1<S-5$　(B)$S-5\le S_1<S$ (C)$S_1<S$　(D)$S<S_1\le S+5$ (E)$S+5<S_1$ (89 自)

(8) 某校高三甲乙丙三班各有 50 位同學，數學模擬考成績的以下累積數折線圖如下(各組不含上限)：



根據上圖中的資料，選出下列正確的選項：
(A)各班成績的中位數，甲班最高 (B)各班的及格人數，丙班最多(60 分(含)以上及格) (C)各班 80 分(含)以上的人數，乙班最多 (D)各班的平均成績，丙班最差 (E)此次模擬考最高分，出現在乙班。　　(89 社)

(9) 甲，乙，丙三位同學參加推薦甄選學科能力測驗，五科的成績如下表所示．設 $S_{甲}$，$S_{乙}$，$S_{丙}$分別代表甲，乙，丙三位同學五科的成績的標準差．請仔細觀察表中數據，判斷下列哪一選項表示 $S_{甲}$，$S_{乙}$，$S_{丙}$的大小關係?

科目／學生	社會	國文	自然	英文	數學
甲	100	70	80	60	50
乙	90	60	70	50	40
丙	80	56	64	48	40

(A)$S_{甲}>S_{丙}>S_{乙}$ (B)$S_{丙}>S_{甲}=S_{乙}$ (C)$S_{甲}>S_{丙}=S_{乙}$ (D)$S_{乙}>S_{甲}=S_{丙}$
(E)$S_{甲}=S_{乙}>S_{丙}$

(10) 有一筆統計資料，共有 11 個數據如下(不完全依大小排列)：2, 4, 4, 5, 5, 6, 7, 8, 11, x 和 y,已知這些數據的算術平均數和中位數都是 6，且 x 小於 y。請選出正確的選項。____________(92 指定甲)

(1) $x+y=14$　(2) $y<9$　(3) $y>8$　(4)標準差至少是 3

(11) 某次數學測驗分為選擇題與非選擇題兩部分。下列的散佈圖中每個點(X,Y)分別代表一位學生於此兩部分的得分， 其中 X 表該生選擇題的得分，Y 表該生非選擇題的得分。

設 Z= X+Y 為各生在該測驗的總分。

共有 11 位學生的得分數據。

試問以下哪些選項是正確的？

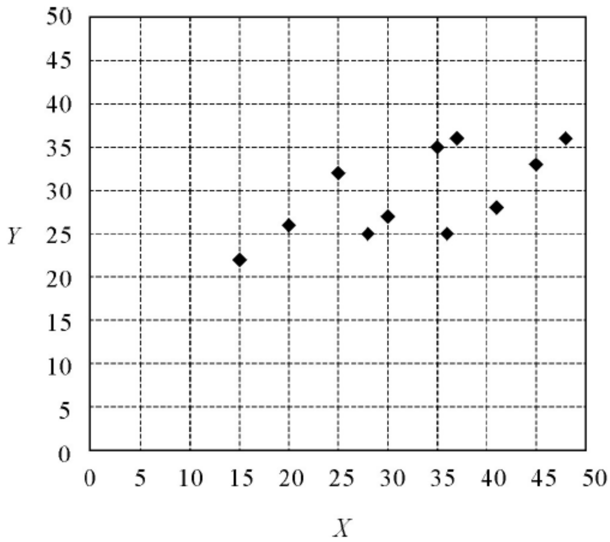
(1) X 的中位數 $Y>$ 的中位數

(2) X 的標準差 $Y>$ 的標準差

(3) X 的全距 $Y>$ 的全距

(4) Z 的中位數= X 的中位數+ Y 的中位數

(95 指定乙)



(12) 某校高一第一次段考數學成績不理想，多數同學成績偏低；考慮到可能是同學們適應不良所致，數學老師決定將每人的原始成績取平方根後再乘以 10 作為正式紀錄的成績。 今隨機抽選 100 位同學，發現調整後的成績其平均為 65 分，標準差為 15 分；試問這 100 位同學未調整前的成績之平均 M 介於哪兩個連續正整數之間？樣本標準差 $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{X})^2}=\sqrt{\frac{1}{n-1}((\sum_{i=1}^{n}x_i^{\,2})-n\overline{X}^{2})}$

(1) $40\le M<41$　(2) $41\le M<42$　(3) $42\le M<43$

(4) $43\le M<44$　(5) $44\le M<45$　(2005 學科)

(13) 某校某班學生的英文測驗的成績結果如下表：

分數	20～30	30～40	40～50	50～60	60～70	70～80	80～90	90～100
人數	1	2	7	10	15	12	6	2

(a)試求其算術平均數 $\overline{x}$ ＝________，標準差 S＝_________。

(b)試求介於 $\overline{x}$ －S 與 $\overline{x}$ ＋S 中的人數占總人數的百分比為_________。

(14) 有 1000 位學生某次月考數學成績之中位數為 72 分，四分位差 12 分，則此次考試數學及格之學生至少有______人。

(15) 某班有 40 人參加考試，老師計算成績後得全班之平均分數為 51 分，標準差為 2 分，但教務處通知：考生阿牛因作弊，其原分數 40 分應改為 0 分。則這班同學考試成績之真正標準差應為__________分。

(16) 設某班有 45 人，這學期兩次期中考的數學成績，每位同學第二次都比第一次多 5 分，下列各數值恆不變的有？
(A)算術平均數(B)第 3 四分位數(C)全距(D)四分位距(E)標準差。

(17) 甲乙兩班段考成績統計結果如下：
甲班 54 人，平均 62 分，母體標準差 6 分乙班 46 人，平均 70 分，母體標準差 5 分，請問兩班學生合併計算，平均成績為多少分？母體標準差為多少分？

(18) 某大公司抽樣調查其名下各分公司員工的月薪，得平均數為 30000 元，標準差為 4000 元。為了激勵員工士氣，董事會提出兩個調薪方案：
甲方案：每人加薪 5000 元。乙方案：每人加薪 5%
求兩方案下員工每月薪水的算術平均數與標準差。

(19) 為調查城鄉差距對中小學英文成績的影響，分別依學生比例抽樣得的結果如下：
自城市抽樣 30 人，其平均 70 分，標準差 5 分
自鄉鎮抽樣 20 人，其平均 60 分，標準差 10 分
若將兩組學生共 50 人的成績合併計算，下列選項哪些是正確的？
(A)合併後的平均低於 70 分
(B)合併後的平均高於 70 分
(C)合併後的標準差高於 6 分
(D)合併後的標準差低於 8 分

(20) 下表是 96 學年度指定考科依全體考生成績計算公佈的數學科成績：
試計算數學甲與數學乙的中位數與四分位距。

科目	頂標	前標	均標	後標	底標
數學甲	62	49	33	20	11
數學乙	72	60	43	27	17
※以上五項標準均取為整數(小數只捨不入)，且其計算均不含缺考生之成績，計算方式如下：					
頂標:成績位於第 88 百分位數之考生成績					
前標:成績位於第 75 百分位數之考生成績					
均標:成績位於第 50 百分位數之考生成績					
後標:成績位於第 25 百分位數之考生成績					
底標:成績位於第 12 百分位數之考生成績					

(21) 九位同學參加數學抽測，其分數如下：
(a)求此 9 個分數的平均數與中位數。

代號	A	B	C	D	E	F	G	H	I
得分	30	40	60	50	70	80	60	90	60

(b)隨機抽樣 3 人，求抽到 3 人的分數的中位數等於母體中位數的機率。
(c)隨機抽樣 3 人，求抽到 3 人的分數的平均數等於母體平均數的機率。

綜合練習解答

(1) 40
(2) 79
(3) (C)
(4) **(1)**
(5) (A)(B)(C)(E)
(6) (C)(D)
(7) (B)
(8) (A)(C)(D)(E)
(9) **(4)(5)**
(10) **(1)(2)**
(11) **(1)(2)(3)**
(12) (5)

設 100 位學生的原始成績爲 x_i，對應的新成績爲 $y_i = 10\cdot\sqrt{x_i}$，

$$\Rightarrow \sum_{i=1}^{100} y_i^{\,2} = (n-1)\cdot s^2 + n\cdot \overline{y}^2 = 99\times 15^2 + 100\times 65^2 = 444775$$

$$\Rightarrow M = \frac{1}{100}\sum_{i=1}^{100} x_i = \frac{1}{100}\sum_{i=1}^{100}\frac{y_i^{\,2}}{100} = \frac{1}{10000}\times 444775 = 44.4775$$

(13) (a)64.3；15.2 (b)67.27%
(14) 750
(15) $\sqrt{65}$=8.06 分
(16) (C)(D)(E)
(17) 65.68 分，6.8333 分
(18) 甲方案：算術平均數 35000 元、標準差 4000 元；
乙方案：算術平均數 31500 元、標準差 4200 元
(19) (A)(C)
(20) 數學甲：中位數 33、四分位距=49−20=29；數學乙：中位數 43、
四分位距=60−27=33。
(21) (a)中位數=60、平均數=60 (b)$\frac{23}{42}$ (c)$\frac{1}{7}$