§3–6 統計量

所謂的「統計量」就是由一組樣本所算出的單一數值，這裡所要討論的統計量有下列兩種型式：表達資料集中情形的統計量、表達資料分散程度的統計量。
(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。

(甲)表達資料集中狀況的統計量

表達資料集中情形的統計量有
算術平均數、加權平均數、幾何平均數、中位數、眾數等。
在介紹這些統計量的計算與意義時，通常將資料分成**已分組資料**或**未分組資料**兩種來介紹。
(1)算術平均數：
一群數值的算術平均數，爲此各項數值的總和除以此群數值的個數所得的商。
(a)未分組資料，求算術平均數：

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

(b)已分組資料，求算術平均數：

變量 X(組中點)	x_1	x_2	……	x_n	總計
次數 f	f_1	f_2	……	f_n	n
變量 $Y=\frac{1}{h}(X-A)$	y_1	y_2	……	y_n	n

一般的算法：
算術平均數$\overline{x}=\frac{1}{n}(f_1x_1+f_2x_2+\ldots+f_nx_n)=\frac{1}{n}\sum_{i=1}^{n} f_ix_i$ 。
縮小組距平移變量：
爲了簡化計算，通常將資料X轉換成資料Y，而$Y=\frac{1}{h}(X-A)$，h爲組距，A爲某組的組中點，這樣一來，Y的數值會變得比較容易計算，先算出$\overline{y}$，就可計算$\overline{x}$，其中$\overline{y}=\frac{1}{h}(\overline{x}-A)$。
[原理]：
若資料X中的每筆數值，都乘以A再加上B，形成資料Y。
即若資料$Y=AX+B$，則$\overline{y}=A\overline{x}+B$。

(c)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爲零。
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。

[例題1] 某校高二丙班 50 名同學之體重分配表如下：

體重（公斤）	人數	組中點
35～40	4	37.5
40～45	3	42.5
45～50	20	47.5
50～55	18	52.5
55～60	2	57.5
60～65	3	62.5

則體重之算術平均數為______________公斤。Ans：49.5 公斤

[例題2] (算術平均數的特點)

若$x_1,x_2,\ldots,x_n$之算術平均數為$\overline{x}$，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$\overline{y}$，證明：

(1) $\sum_{i=1}^{n} x_i = n\overline{x}$ (2) $\sum_{i=1}^{n}(x_i-\overline{x})=0$ (3) $\overline{y}=A\overline{x}+B$。

(練習1) 100 戶人口每月收入之次數分配表如下：

組別	30 ~ 40	40 ~ 50	50 ~ 60	60 ~ 70	70 ~ 80	80 ~ 90	90 ~ 100	100 ~ 110
人數	5	11	20	33	18	8	3	2

求其算術平均數。　Ans：64.4

(練習2) 某班段考的數學成績經統計後，得到平均分數為 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，平均分數為__________分。Ans：82

(練習3) 設有 51 個數值之算術平均數為 60，後來發覺其中「85」一數必須剔除，試求剔除「85」一數後所餘 50 個數值之算術平均數。Ans：59.5

(練習4) 某一組資料為 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為(A) $\frac{\overline{y_1}+\overline{y_2}}{2}$　(B) $\frac{\overline{y_1}+\overline{y_2}}{N_1+N_2}$　(C) $\frac{N_1\overline{y_1}+N_2\overline{y_2}}{N_1+N_2}$

(D) $\frac{N_1\overline{y_1}+N_2\overline{y_2}}{2}$ 。Ans：(C)

(2)加權平均數：

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

W定義為W=$\sum_{i=1}^{n} x_i w_i$ 。

(a)為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為學生的學期總平均成績，顯然不適當，於是為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為何？

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。

[例題3] 小安第二次月考之成績如下表，若以上課時數來做為權數，
請求出其加權平均數。Ans：73.62

科目	國文	英文	數學	物理	化學
上課時數	6	4	5	3	3
成績	78	75	65	71	80

(練習5) 某次月考某生各科成績及每週上課時數如下表：

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

若該生的加權平均介於 60～65 分之間，則數學成績可能是幾分
(A) 45　(B) 49　(C) 55　(D) 70　(E) 74　分。Ans：(B)(C)(D)

(3)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$\dfrac{62}{5}$。

(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= $\dfrac{20\%-20\%}{2}$ =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。

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

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) 大明開設一公司，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%

(4)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 $\frac{n+1}{2}$ 個數。

n 爲偶數時，中位數是指按大小順序排列後之第 $\frac{n}{2}$個數與第 $\frac{n}{2}+1$ 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 ……… L_{i-1}~U_{i-1} L_i~U_i …….. L_k~U_k	f_1 ………. f_{i-1} f_i …….. f_k	$C_1=f_1$ ……… $C_{i-1}=f_1+f_2+f_3+\ldots+f_{i-1}<\frac{n}{2}$ $C_i=f_1+f_2+\ldots+f_{i-1}+f_i\geq\frac{n}{2}$
…總計…	n	

Step2：找出 Me 所在的組別

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

(b)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。

[例題6] 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)甲組之中位數爲__________ (2)乙組之中位數爲__________。

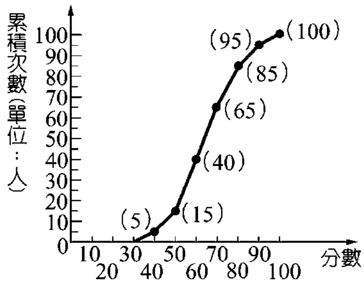
Ans：(1)53 (2)68.5

[例題7] 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	

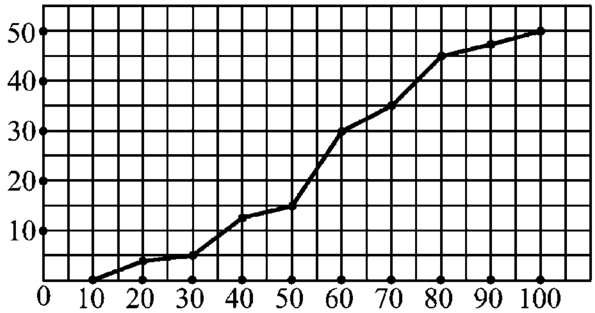
求中位數。Ans：62

(練習8) 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)算術平均數爲__________；
(2)中位數爲__________。
（四捨五入成整數）
Ans：(1)64.5(2)64

(練習9) 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，則成績的中位數爲__________（取到整數，小數點以下四捨五入）



Ans：57

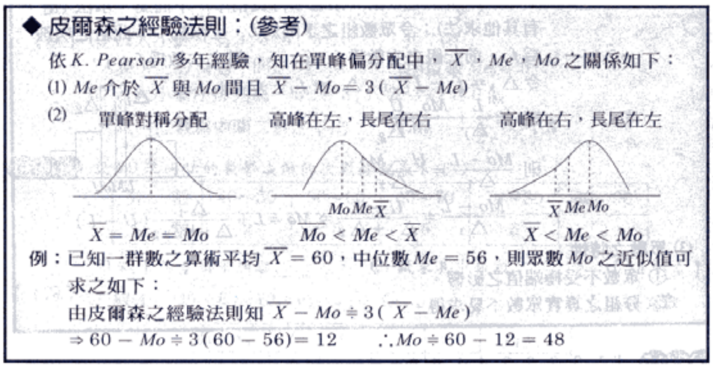
(練習10) 12 位同學在罰球線上投籃 10 次，投中次數分別爲 3,2,3,7,5,3,6,4,1,3,6,8
(1)試求這 12 位同學投中次數的中位數。
(2)若小安投中 6 次，請問這 13 位同學投中次數的中位數。
Ans：(1)3.5 次 (2)4 次

(5)眾數：
在一組數值資料中，出現次數最多的數值稱爲「眾數」。

[例題8] 某班前 18 名同學期末考數學成績如下：
55,60,65,45,70,75,90,85,80,75,55,60,45,70,85,75,75,60
求其眾數。 Ans：75

(練習11) 25 個學生參加投籃遊戲，每個人投籃 10 次，他們的進球數依序爲 7，4，6，2，2，7，4，4，3，1，5，4，5，9，4，1，4，9，8，5，4，4，1，4，6，則這群學生進球數的眾數爲__________。Ans：4

(6)資料的分布情形：



◆ 皮爾森之經驗法則：(參考)

依 K. Pearson 多年經驗，知在單峰偏分配中，$\overline{X}$，Me，Mo 之關係如下：

(1) Me 介於 $\overline{X}$ 與 Mo 間且 $\overline{X}-Mo \doteq 3(\overline{X}-Me)$

(2) 單峰對稱分配　高峰在左，長尾在右　高峰在右，長尾在左

$\overline{X}=Me=Mo$　$Mo<Me<\overline{X}$　$\overline{X}<Me<Mo$

例：已知一群數之算術平均 $\overline{X}=60$，中位數 $Me=56$，則眾數 Mo 之近似值可求之如下：

由皮爾森之經驗法則知 $\overline{X}-Mo \doteq 3(\overline{X}-Me)$

$\Rightarrow 60-Mo \doteq 3(60-56)=12$　$\therefore Mo \doteq 60-12=48$

（乙）表達資料分散程度的統計量

表達資料分散程度的統計量有

全距、四分位差、標準差等。

介紹這些統計量的計算與意義時，通常將資料分成**已分組資料**或**未分組資料**兩種來介紹。

(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。

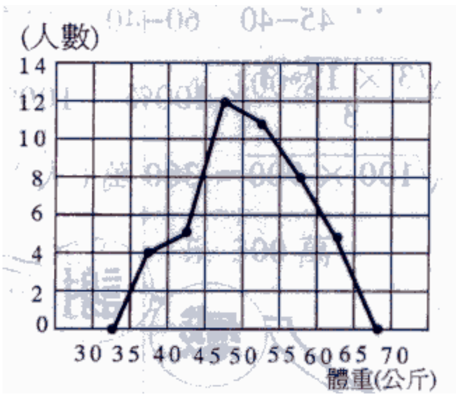
[例題9] 右圖是某班同學的次數分配折線圖，求其全距。

Ans：30

[解法]：

因為 30~35 與 65~70 的人數均為 0，

故 R=65−35=30。



(2)四分位差：

(a)將一群統計資料的數據，　依照其大小順序，由小到大排成一列，設中位數為M_e，在此數列中，位在M_e後段各數的中位數(稱為第 3 個四分位數，以Q_3表示)，又位在M_e前段各數的中位數(稱為第 1 個四分位數，以Q_1表示)，則四分位差Q.D.定義為Q_3-Q_1。

(b)未分組資料求Q.D.：

將n個資料按大小順序排列爲$x_{(1)}\leqq x_{(2)}\leqq x_{(3)}\leqq\cdots\leqq x_{(n)}$，則不論$n$爲 $2m$或 $2m+1$，其中位數以上與以下的個數必等於m個。

(1°)當$m=2k+1$時，$Q_1=x_{(k+1)}$，$Q_3=x_{(n-k)}$，則Q.D.= (Q_3-Q_1)

(2°)當$m=2k$時，$Q_1=\frac{1}{2}(x_{(k)}+x_{(k+1)})$，$Q_3=\frac{1}{2}(x_{(n-k)}+x_{(n-k+1)})$，則Q.D.= (Q_3-Q_1)

(c)已分組資料求Q.D.

Step1：仿造求中位數的方法，求出Q_3及Q_1

$$\frac{Q_1-L_{Q_1}}{U_{Q_1}-L_{Q_1}}=\frac{\frac{n}{4}-C_{Q_1}}{f_{Q_1}}，\frac{Q_3-L_{Q_3}}{U_{Q_3}-L_{Q_3}}=\frac{\frac{3n}{4}-C_{Q_3}}{f_{Q_3}}$$

L_{Q_1}、L_{Q_3}分別代表Q_1、Q_3所在組之下限，U_{Q_1}、U_{Q_3}分別代表Q_1、Q_3所在組之上限；C_{Q_1}、C_{Q_3}代表L_{Q_1}、L_{Q_3}之以下累積次數，f_{Q_1}、f_{Q_3}代表Q_1、Q_3所在組之次數。

Step2：四分位差=Q_3-Q_1

(d)特性：受隨機抽樣不確定的影響較小；感應不靈敏。

(e)注意事項

①當整個統計資料均勻分散時，使用全距表示資料分散情形較爲簡單明確。

②當整個統計資料具有集中趨向時，使用四分位差表示資料分散情形較爲正確可靠。

[例題10] 甲，乙兩班分別有 13 個學生與 12 個學生參加學力測驗，兩班學生的成績分別爲

甲班：8，7，18，15，18，21，18，14，15，16，20，20，21。

乙班：17，17，15，16，8，18，18，19，20，18，19，21。

則甲班成績的四分位差爲______，乙班成績的四分位差爲______。

Ans：5.5；2.5

[例題11] 某班月考的數學成績統計如下：求：

分　數	人　數	分　數	人　數	分　數	人　數
0～10	0	40～50	16	80～90	13
10～20	1	50～60	23	90～100	4
20～30	5	60～70	28	合　計	120
30～40	9	70～80	21		

(1)全距為________；(2)四分位差為__________。
Ans：(1)90(2)24.44

(練習12) 某人觀測一星期中某時刻的氣溫，得攝氏溫度如下：
15，12，13，8，9，11，9，則這星期中，此時刻的溫度記錄之全距為
。Ans：7

(練習13) 設有十一個數據資料為：2，15，16，17，17，18，18，19，19，72，
90，試求：(1)全距為_______；(2)四分位差為________。
Ans：(1)88(2)3

(練習14) 某市 100 戶房客每月所付房租如下表，試求其四分位差為_______。

房租（元）	0～100	100～200	200～300	300～400	400～500
戶　數	1	32	50	15	2

Ans：109

(3)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$\overline{x}$，

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

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$\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})}{n}$即可知道**平均離差**爲何。但不幸的是，

因爲$n\overline{x}=\sum_{i=1}^{n}x_i$，也就是說$\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})}{n}=0$。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**爲$\dfrac{\sum_{i=1}^{n}|x_i-\overline{x}|}{n}$。

例如：

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

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

乙的平均絕對離差爲$\frac{1}{3}(|-1|+|1|+|0|)=\frac{2}{3}$。

$\Rightarrow$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，即爲變異數的概念。

(3)變異數

定義：

一組資料$x_1,x_2,x_3,\ldots\ldots,x_n$的**母體變異數**$\sigma^2$爲$\dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$。$\mu$爲母體的算術平均數

一組資料$x_1,x_2,x_3,\ldots\ldots,x_n$的**樣本變異數**$s^2$爲$\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}$。$\overline{x}$爲樣本的算術平均數

由於變異數的單位是資料單位的平方，(如本來單位是cm，但變異數爲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**。

(4)標準差：

(a)未分組的標準差公式

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

樣本標準差s=$\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2}=\sqrt{\frac{1}{n-1}[\sum_{i=1}^{n}x_i^2-n(\overline{x})^2]}$

[例題12] 小安第一次月考六科的平均成績為 80 分，若已知其中五科的成績為 68,80,80,80,86，則其成績的標準差為_______分。 Ans：6 分

[例題13] 有 10 個資料 x_1，x_2，……，x_{10}，已知 $\sum_{i=1}^{10}x_i=155$，$\sum_{i=1}^{10}x_i^{\ 2}=2551$，求其算術平均數為___________，標準差為_________。

Ans：15.5；3.9

(b)已分組的標準差公式：

變量 X(組中點)	x_1	x_2	……	x_n	總計
次數 f	f_1	f_2	……	f_n	n

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

樣本標準差σ=$\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}f_i(x_i-\overline{x})^2}=\sqrt{\frac{1}{n-1}[\sum_{i=1}^{n}f_ix_i^2-n(\overline{x})^2]}$

(c)標準差的簡化計算(縮小組距平移變量)：

變量 X(組中點)	x_1	x_2	……	x_n	總計
次數 f	f_1	f_2	……	f_n	n
變量 $Y=\frac{1}{h}(X-A)$	y_1	y_2	……	y_n	n

爲了簡化計算，通常我們會將資料X轉換成資料Y，而$Y=\frac{1}{h}(X-A)$，h爲組距，A爲某組的組中點，這樣一來，Y的數值會變得比較容易計算，先算出Y的標準差S_Y，就可計算X的標準差S_X。

[原理]：

若資料X中的每筆數值，都乘以A再加上B，形成資料Y即資料Y=AX+B，

則$S_Y=|A|S_X$。

[證明]：

(5)標準差的特性：

(a)以算術平均數爲中心的標準差，較任何其它的平均數爲中心的標準差小。

(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。

(6)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=aX+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、四分位差、標準差)=$|a|$(X 的差異量數)

[例題14] 某次競試 100 人參加，考試結果如下表，試求

成績	20 ~ 30	30 ~ 40	40 ~ 50	50 ~ 60	60 ~ 70	70 ~ 80	80 ~ 90
人數	3	6	10	16	30	21	14

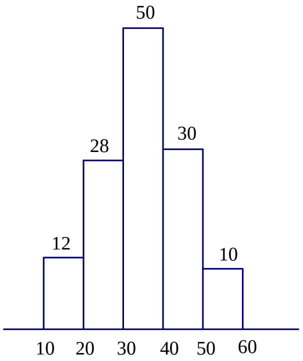
(1)算術平均數 (2)標準差。Ans：(1)63.3 (2)15.2

[解法]：

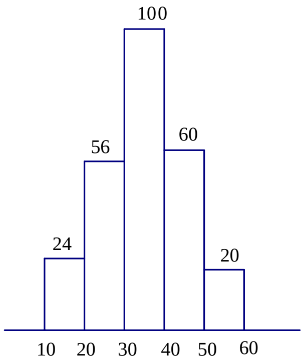
X(成績組中點)	25	35	45	55	65	75	85
$Y=\dfrac{1}{10}(X-65)$							
人數	3	6	10	16	30	21	14

[例題15] 下列五個直方圖表示的資料，何者之標準差最大？ (2005 指定乙)

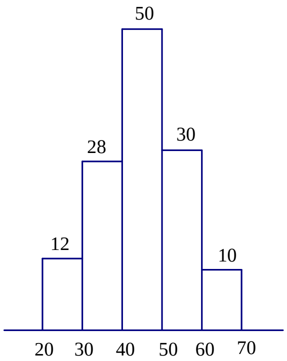
(1)



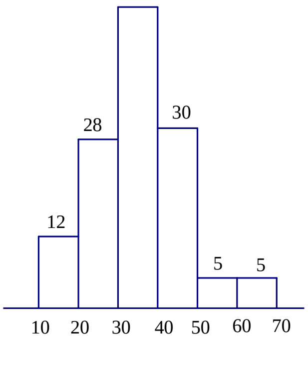
(2)



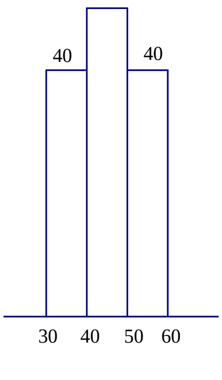
(3)



(4)



(5)



Ans：(4)

[例題16] 測量一物件 9 次，得其長度(公尺)爲 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 得一組新的數距爲 3，6，1，5，4，8，6，7，5，問下列選項何者爲真？ (A)新數據的算術平均數爲 5 (B)新數據的標準差爲 2 (C)原數據的算術平均數爲 2.45 (D)原數據的標準差爲 0.2 (E)原數據的中位數爲 2.45

Ans：(A)(B)(C)(E) (88 學科)

[例題17] (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】：$\overline{X}=\dfrac{20\times 75+30\times 60}{20+30}=\dfrac{2\times 75+3\times 60}{2+3}=66$

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

$\Rightarrow \sigma^2=\dfrac{20(25+81)+30(36+36)}{50}=\dfrac{2\times 106+3\times 72}{5}=85.6$

$\therefore \sigma=\sqrt{85.6}\doteqdot 9.3$

【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 $\overline{X}=75$，$S_1=5.0$

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

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

則 $\overline{Y}=60$，$\sigma_2=6.0$

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

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

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

標準差 $\sigma=\sqrt{\dfrac{222080}{50}-66^2}=\sqrt{85.6}\doteqdot 9.25\doteqdot 9.3$

【註】：

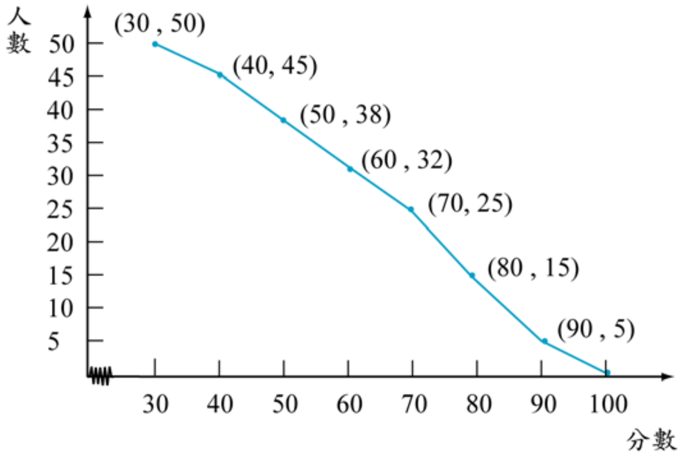
本題之標準差如果改爲樣本標準差 $S=\sqrt{\dfrac{\Sigma(x_i-\overline{X})^2}{n-1}}$

則 $S^2=\dfrac{19\cdot 5^2+20(75-66)^2+29\cdot 6^2+30(60-66)^2}{20+30-1}$

$=\dfrac{475+1620+1044+1080}{49}=\dfrac{4219}{49}\doteqdot 86.1$

$\therefore S=\sqrt{86.1}=9.28\doteqdot 9.3$

(練習15) 某班段考成績之累積次數分配曲線圖如下：（採相同組距 10，且不含上限）
試求 (1)全距 (2)四分位差 (3)標準差 .



Ans：(1)70(2)32(3)18.55

(練習16) 求下列各數的標準差：2,3,7,8,10　　Ans：$\sqrt{9.2}$

(練習17) 求 1700,1800,1900,2000,2100,2200 之標準差。　Ans：$\dfrac{50\sqrt{105}}{3}$

(練習18) 高二某班第二次月考數學成績統計如下表，試求下列各值：
（四捨五入取到小數第一位）

分數	20～30	30～40	40～50	50～60	60～70	70～80	80～90	90～100
人數	1	2	3	4	10	15	10	5

(1)算術平均數 $\overline{x}$ =＿＿＿＿＿；(2)中位數 Me＝＿＿＿＿＿；
(3)四分位差 Q.D.＝＿＿＿＿＿＿；(4)標準差 S＝＿＿＿＿＿＿。
Ans：(1)71(2)73.3(2)20(4)16.4

(練習19) 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

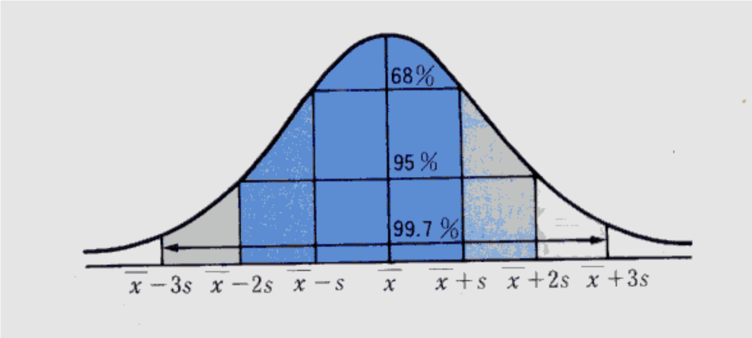
(練習20) 有一群資料X：1，1，2，3，5，5，7，8，9，9，另一群資料Y：2001，2001，2002，2003，2005，2005，2007，2008，2009，2009，則下列何者正確？
(A) X之中位數＝5　(B) Y之算術平均數＝2005
(C) Y之中位數＝2005　(D) X之標準差$S_x=3$　(E) Y之標準差$S_y=S_x$。
Ans：(A)(B)(C)(D)(E)

（丙）常態分布

(1)常態分布：

平均數與標準差是資料分析最重要的兩個統計量，對於常態分布的資料，我們由次數分配呈鐘形知道中間部分佔的比例較大，愈往兩旁所佔的比例愈小，但比例大約是多少呢？底下，讓我們對常態分布的資料估計這些比例。

當一組資料的直方圖呈鐘形(即資料是常態分布)，而且也知道此組資料的平均數 $\overline{x}$，標準差 s，就能估算出大約有 68%的資料落在區間($\overline{x}-s$, $\overline{x}+s$)內，有 95%的資料落在區間($\overline{x}-2s$, $\overline{x}+2s$)內，有 99.7%的資料落在區間($\overline{x}-3s$, $\overline{x}+3s$)內。



(2)標準計分：

標準分數 (standard score)= $\dfrac{\text{觀測值}-\text{算術平均數}}{\text{標準差}}$ 。

標準分數爲 1 的意義是說：觀測值在算數平均數之上 1 個標準差的位置；標準分數爲−2 的意義是說：觀測值在算數平均數之下 2 個標準差的位置。標準分數可以用來比較不同分布中的值。在使用時，分布必須至少是大致對稱的，標準分數才適用。

[例題18] 某校有學生 1000 位，數學段考成績爲常態分布，平均成績爲 70 分，標準差爲 10 分，則下列敘述何者正確？
(A)數學不及格人數大約有 160 位。
(B)超過 90 分的人數大約有 25 位。
(C)某生成績爲 80 分，其全校排名約 160 名。
(D)將每位學生成績減 70 分，再除以 10 分，則此新成績的平均數爲 0 分。
(E)將每位學生成績減 70 分，再除以 10 分，則此新成績的標準差爲 1 分。
Ans：(A)(B)(C)(D)(E)

[例題19] Jane 在 SAT 的語言部份得了 600 分，她的朋友 Joy 則是參加 ACT 測驗，在語言部份拿了 21 分。根據測驗單位的統計結果，SAT 與 ACT 大致上呈常態分布，且它們的算術平均數與標準差分別爲 500，110；18，6，請問誰的標準計分比較高。 Ans：Jane

(練習21) 某校有學生 1000 位，數學段考成績呈常態分布，平均成績 70 分，標準差 10 分，請概估：
(1)此次數學段考不及格的學生約有幾位？
(2)成績超過 90 分的有幾位？
(3)某生成績 80 分，他在全校大約排第幾名？
Ans：(1)160 位　(2)25 位　(3)第 160 名

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為_____%。(90 學科)

(2) 某校學生有 1000 人，數學段考成績為常態分配，平均成績為 65.24 分，標準差 5.24 分，試問全校約有多少人數學成績低於 60 分？
(A)80 (B)160 (C)240 (D)320 (E)400 人　(91 學科)

(3) 在某項才藝競賽中，為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為此參賽者的比賽成績。現在有一位參賽者所獲 15 位評審的平均成績為 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為________分。(96 學科)

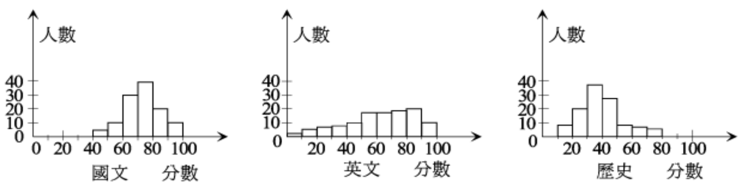
(4) 某校想要了解全校同學是否知道中央政府五院院長姓名，出了一份考卷。該卷共有五個單選題，滿分 100 分，每題答對得 20 分，答錯得零分，不倒扣。閱卷完畢後，校方公佈每題的答對率如下：

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)

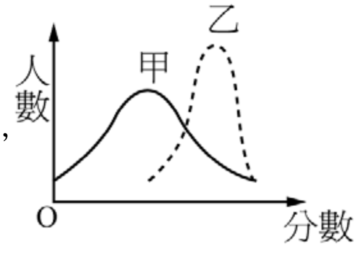
(5) 下列 5 組資料(每組各有 10 筆)
A:1,1,1,1,1,10,10,10,10,10
B:1,1,1,1,1,5,5,5,5,5
C:4,4,4,5,5,5,5,6,6,6
D:1,1,2,2,3,3,4,4,5,5
E:1,2,3,4,5,6,7,8,9,10
試問哪一組資料的標準差最大？
(1)A (2)B (3)C (4)D (5)E　(89 學科)

(6)



上列三圖為某年級國文，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。

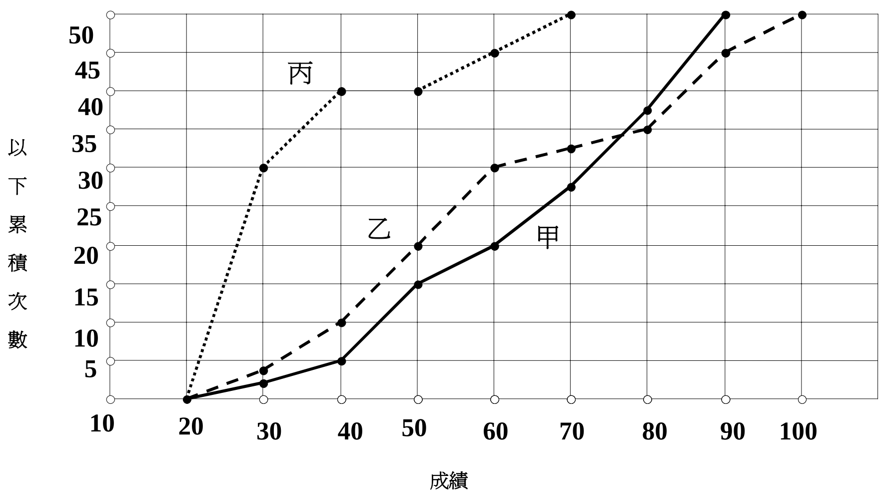
(7) 某年聯考甲、乙兩科成績的直方圖如圖所示，
(由於考生人數眾多，成績分布的直方圖可視為平滑的曲線)，
則下列哪些敘述是正確的？
(A)甲的算術平均數比乙的算術平均數大



(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

(8) 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$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2}$，而$\bar{x}=\frac{1}{n}\cdot\sum_{i=1}^{n}x_i$ (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)

(9) 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)

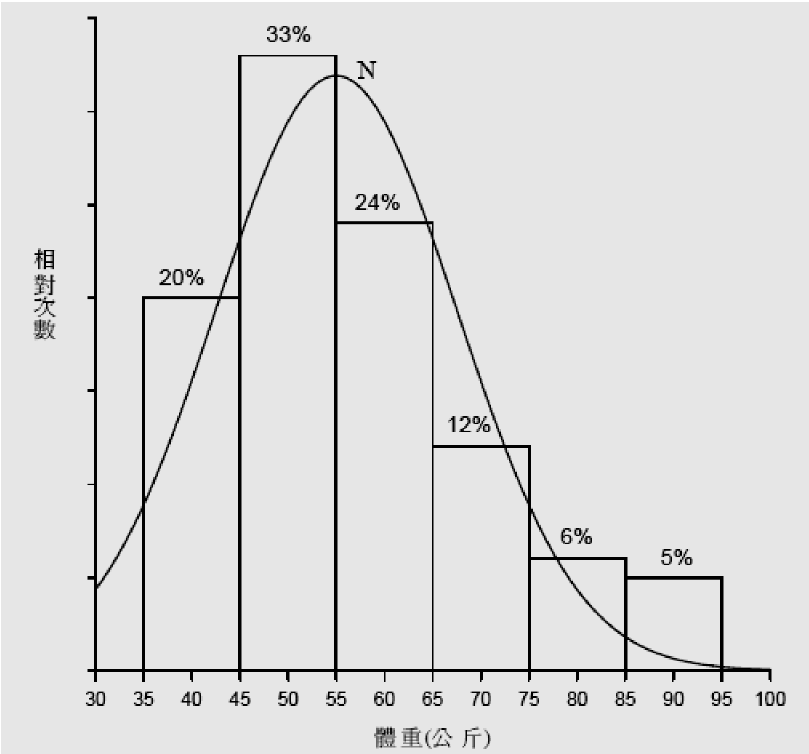
(10) 甲，乙，丙三位同學參加推薦甄選學科能力測驗，五科的成績如下表所示．設$S_甲$，$S_乙$，$S_丙$分別代表甲，乙，丙三位同學五科的成績的標準差．請仔細觀察表中數據，判斷下列哪一選項表示$S_甲$，$S_乙$，$S_丙$的大小關係?

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)$S_甲=S_乙>S_丙$

(11) 下圖是根據 100 名婦女的體重所作出的直方圖（圖中百分比數字代表各體重區間的相對次數，其中各區間不包含左端點而包含右端點）。該 100 名婦女體重

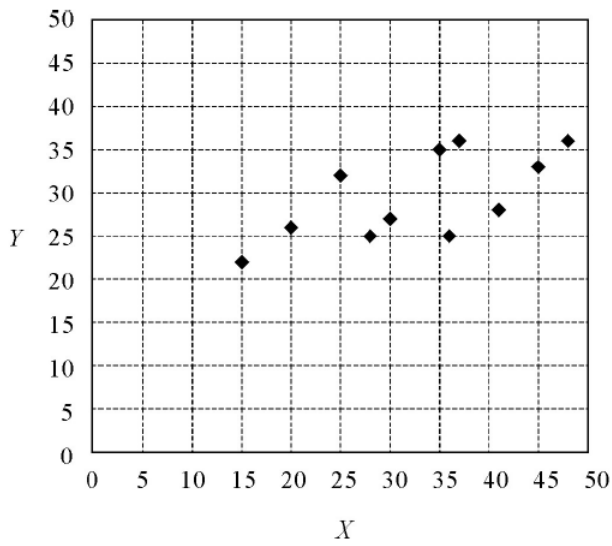
的平均數爲 55 公斤，標準差爲 12.5 公斤。曲線 N 代表一常態分佈，其平均數與標準差與樣本值相同。在此樣本中，若定義「體重過重」的標準爲體重超過樣本平均數 2 個標準差以上（即體重超過 80 公斤以上），則下列敘述哪些正確?



(1) 曲線 N（常態分佈）中，在 55 公斤以上所佔的比例約爲 50%。
(2) 曲線 N（常態分佈）中，在 80 公斤以上所佔的比例約爲 2.5%。
(3) 該樣本中，體重的中位數大於 55 公斤。
(4) 該樣本中，體重的第一四分位數大於 45 公斤。
(5) 該樣本中，「體重過重」（體重超過 80 公斤以上）的比例大於或等於 5%。
(2006 學科能力測驗)

(12) 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

(13) 某次數學測驗分爲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 爲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)



(14) 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爲_________。

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

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

(17) 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)四分位差(E)標準差。

(18) 高二 A 班學生 48 人，工藝學期成績經統計後得知：①中位數爲 68 分；②四分位差爲 8 分，由①，②可以斷言：該班至少有 36 人的工藝成績超過 60 分，試說明它的理由。

(19) 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，平均成績爲多少分？母體標準差爲多少分？

綜合練習解答

(1) 40

(2) (B)

(3) 79

(4) (C)

(5) **(1)**

(6) (A)(B)(C)(E)

(7) (C)(D)

(8) (B)

(9) (A)(C)(D)(E)

(10) **(4)(5)**

(11) (E)

(12) **(1)(2)**

(13) **(1)(2)(3)**

(14) (a)64.3；15.2 (b)67.27%

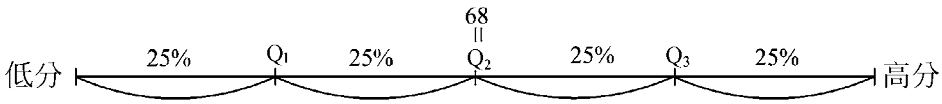
(15) 750

(16) $\sqrt{65}=8.06$ 分

(17) (C)(D)(E)

(18) 因$Q_2-Q_1\leqq Q_3-Q_1\Rightarrow 68-Q_1\leqq 8$（$\because$Q.D.$=Q_3-Q_1=8$）$\Rightarrow Q_1\geqq 60$ 工藝成績≧

Q_1 之學生人數至少爲 $48\times\frac{75}{100}=48\times\frac{3}{4}=36$（人）



(19) 65.68 分，6.8333 分