第十六章　　統計初步

在日常生活中，我們經常要收集數據、累積資料，然後通過整理、計算與分析，從中找出規律，這就要用到統計方法。本章介紹統計的一些初步知識。

16.1　總體與樣本

看下面的例子：

1. 有關部門要了解某個地區九年級學生的體重，以掌握這些學生的身體發育情況。於是，這個地區所有九年級學生的體重，就是所要考察的對象。但由於這個地區的九年級學生較多，逐一考察比較困難，於是只能抽查其中一部份(比如說，2000 名)學生的體重，然後根據這一部分學生的體重去估計所有學生的體重情況。

2. 要了解一塊玉米地裡所有單株玉米的情況(例如，了解它們的平均產量)。這樣，這塊地裡所有單株玉米的產量就是所要考察的對象。但是，由於這塊地裡的玉米株數很多，無法一一加以考察，只能從中抽取一部份(比如說，抽取 100 株)單株玉米，然後用這一部分單株玉米的產量，去估計這塊地裡所有單株玉米的產量。

3. 要考察某批砲彈的殺傷半徑(例如，考察殺傷半徑的範圍)。這種考察是帶有破壞性的，因為測量一個就得發射一個。這樣，即使砲彈的個數不是很多，也不可能一一進行試驗，而是從中抽取一部份(比如說，抽取 10 發)砲彈來進行試驗，然後用這一部分砲彈的殺傷半徑，去估計這批砲彈中所有砲彈的殺傷半徑。

我們所要考察的對象之全體叫做**總體**，其中每一個考察對象叫做**個體**，從總體中抽取的一部份個體叫做總體的一個**樣本**，樣本中個體的數目叫做**樣本的容量**。

在第一個例子中，所指地區九年級學生體重的全體是總體，每個學生體重是個體，從中抽取的 200 名學生之體重是總體的一個樣本，樣本的容量是 200。

在第二個例子中，玉米地裡單株玉米產量的全體是總體，每個單株玉米的產量是個體，從中抽取的 100 個單株玉米之產量是總體的一個樣本，樣本的容量是 100。

在第三個例子中，砲彈殺傷半徑的全體是總體，每發砲彈的殺傷半徑是個體，從中抽取的 10 發砲彈是總體之一個樣本，樣本的容量是 10。

我們看到，在統計裡所談到的考察對象，是一種數量指示。如在第一個例子中，我們考察的對象是所指地區各九年級學生的體重，而不是籠統指這些學生本身。

總體中包含的個體數往往很多(如第一個例子與第二個例子)，有時總體中包含的個數不是很多，但考察時帶有破壞性(如第三個例子)，因此，我們通常是從總體中抽取一個樣本，然後根據樣本的特性去估計總體的相應特性。

練　習

1. 說明在以下問題中，總體、個體、樣本、樣本的容量各指什麼。
   (1) 為了考察某地國中畢業生數學升學考試的情況，從中抽查了 200 名考生的成績；
   (2) 為了了解一批燈泡的使用壽命，從中抽取了 10 個進行試驗；
   (3) 為了考察某公園一年中每天進園的人數，在其中的 30 天裡對進園的人數進行了統計。
2. 舉一個在實際生活中通過樣本研究總體的例子。

16.2　平均數

某班第一小組在一次數學測驗中的成績如下：

86、91、100、72、93、89、90、85、75、95。

這個小組的平均成績是多少？

很明顯，這個平均成績就是

$$\frac{86+91+100+72+93+89+90+85+75+95}{10}=87.6 \text{。}$$

一般地，如果有 n 個數

$$x_1 \text{、} x_2 \text{、} \cdots \text{、} x_n \text{，}$$

那麼

$$\bar{x}=\frac{1}{n}(x_1+x_2+\cdots+x_n) \qquad (1)$$

叫做這 n 個數的**平均數**，$\bar{x}$ 讀作「x-Bar」。

為了書寫方便，有時將 $x_1+x_2+\cdots+x_n$ 記作 $\sum_{i=1}^{n} x_i$ ，其中「Σ」是求和符號，讀作「sigma」。$\sum_{i=1}^{n} x_i$ 讀作「$\Sigma-x-i$，i 從 1 到 n」，表示從 x_1 加到 x_n 的和。於是公式(1)可以寫成

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

練 習

1. 計算下列各組數的平均數：
   (1) 1.4、2.8、5.3、3.7、2.3；
   (2) −0.3 、0.2、0.3、 −0.4 、0.6、0.3、0.5、 −0.4 。
2. 用求和符號「Σ」表示下列各式：。
   (1) $P_1+P_2+\cdots+P_{10}$ ； (2) $y_1+y_2+\cdots+y_n$ 。

在一次畢業考中，考生有一萬多名。我們想了解這一萬多名考生的數學平均成績。但如果將他們的成績全部加在一起再除以考生總數，十分麻煩。這時，可以採用樣本估計總體的方法，即從中抽查部分考生的成績，用他們的平均成績去估計所有考生的平均成績。

總體中所有個體的平均數叫做**總體平均數**，樣本中所有個體的平均數叫做**樣本平均數**。對於由一萬多名考生的成績組成之總

體來說，所有考生的平均成績就是總體平均數，所抽查的部分考生之平均成績就是樣本平均數。通常我們用樣本平均數去估計總體平均數。一般來說，樣本容量越大，這種估計也就越精確。例如，抽查的考生越多，所抽查的考生之平均成績就越接近總平均成績。

【例 1】 從參加畢業考試的學生中，抽查了 30 名學生的數學成績，分數如下：

90	84	84	86	87	98	78	82	90	93
68	95	84	71	78	61	94	88	77	100
70	97	85	68	99	88	85	92	93	97

計算樣本平均數(結果保留到個位)。

解 $\bar{x}=\frac{1}{30}(90+84+\cdots+97)=\frac{2562}{30}\approx 85$。

即樣本平均數為 85。

於是可以估計，參加畢業考試的學生之數學平均成績約為 85 分。

【例 2】 從一批機器零件毛坯中取出了 20 件，稱它們的重量如下(單位：kg)：

210	208	200	205	202	218	206	214	215	207
195	207	218	192	202	216	185	227	187	215

計算樣本平均數(結果保留到個位)。

解 $\bar{x}=\frac{1}{20}(210+208+\cdots+215)=\frac{4129}{20}\approx 206\,\text{kg}$。

即樣本平均數為 206 kg。

於是可以估計，這批機器零件毛坯平均每件約重 206 kg。

本題中的樣本數據大，而且都在 200 左右波動。這時，也可以採用下面的算法：

將上面各數據同時減去 200，得到一組新數據：

10	8	0	5	2	18	6	14	15	7
–5	7	18	–8	2	16	–15	27	–13	15

計算這組新數據的平均數，得

$$\overline{x'}=\frac{1}{20}(10+8+\cdots+15)=\frac{129}{20}\approx 6\text{ ，}$$

於是，所求的平均數應該是

$$\bar{x}=\overline{x'}+200\approx 6+200=206\,\mathrm{kg}\text{，}$$

所得的結果與前面之結果一樣。我們看到，採用這種算法，可以使參與計算的數據變小，計算起來較簡便。

一般地，如果將一組數據 x_1 、 x_2 、⋯、 x_n 同時減去一個適當的常數 a [6]，得到

$$x_1'=x_1-a\text{ 、 }x_2'=x_2-a\text{ 、⋯、 }x_n'=x_n-a\text{ ，}$$

那麼

$$x_1=x_1'+a\text{ 、 }x_2=x_2'+a\text{ 、⋯、 }x_n=x_n'+a\text{ ，}$$

因此，

$$\begin{aligned}
\bar{x}&=\frac{1}{n}(x_1+x_2+\cdots+x_n)\\
&=\frac{1}{n}[(x_1'+a)+(x_2'+a)+\cdots+(x_n'+a)]\\
&=\frac{1}{n}[(x_1'+x_2'+\cdots+x_n')+na]\\
&=\frac{1}{n}(x'+x_2'+\cdots+x_n')+\frac{1}{n}\bullet na\\
&=\overline{x'}+a
\end{aligned}$$

即

$$\boxed{\bar{\boldsymbol{x}}=\overline{\boldsymbol{x}'}+\boldsymbol{a}} \qquad (2)$$

在例 2 的第二種算法中，正是利用了公式(2)。

[6] 常數 a 通常取接近於樣本平均數(約略估計)的較「整」的數。

【例 3】某工人在 30 天中加工一種零件的日產量，有 2 天是 51 件、3 天是 52 件、6 天是 53 件、8 天是 54 件、7 天是 55 件、3 天是 56 件、1 天是 57 件。計算這個工人 30 天中的平均日產量(結果保留到個位)。

解　在上面 30 個數據中，51 出現 2 次、52 出現 3 次、53 出現 6 次、54 出現 8 次、55 出現 7 次、56 出現 3 次、57 出現 1 次。由於這組數據都比 50 稍大一點，我們利用公式(2)計算它們的平均數，並將公式中的常數 a 取作 50。

將數據 51、52、53、54、55、56、57 同時減去 50，得到

1　2　3　4　5　6　7

它們出現的次數依次是

2　3　6　8　7　3　1

那麼，這組新數據的平均數是

$$\overline{x}' = \frac{1\times2+2\times3+3\times6+4\times8+5\times7+6\times3+7\times1}{30} = \frac{118}{30} \approx 4$$

根據公式(2)，

$$\overline{x} = \overline{x}' + a \approx 4+50 = 54 \text{ 件，}$$

即這個工人 30 天中的平均日產量為 54 件。

於是可以估計，這個工人在這 30 天前後的一段時間內，平均日產量約為 54 件。

一般來說，如果在 n 個數中，x_1 出現 f_1 次、x_2 出現 f_2 次、…、x_k 出現 f_k 次(這裡 $f_1+f_2+\cdots+f_k=n$)，那麼根據公式(1)，這 n 個數的平均數可以表示為

$$\boldsymbol{\overline{x} = \frac{x_1 f_1 + x_2 f_2 + \cdots + x_k f_k}{n}}\text{[7]} \qquad (1')$$

或簡記作

[7] 這個平均數叫做**加權平均數**，其中 f_1、f_2、…、f_k 叫做**權**。

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{k}x_i f_i \qquad (1')$$

正如例 3 那樣，當樣本中有不少數據多次重複出現時，利用公式(1′)計算樣本平均數比較簡便。

練　習

1. 抽測了 10 株茶樹的高，結果如下(單位：cm)：

   25　41　40　43　22　14　19　39　21　42

   計算樣本平均數(結果保留到個位)。

2. 利用公式(2)計算下面各組數據的平均數(結果保留到個位)：

   (1) 105、103、101、100、114、108、110、106、98、102；

   (2) 4203、4204、4200、4194、4204、4201、4195、4199；

3. 從一所中學某年級的六個班中抽測了 24 名男同學的身高，結果如下(單位：cm)：

155	157	159	162	162	163	164	164
165	165	165	166	166	167	167	167
168	169	169	170	171	171	172	174

   這個年級男學生的平均身高約是多少(結果保留到個位)？

4. 在一個班中的 40 名學生中，14 歲的有 5 人、15 歲的有 30 人、16 歲的有 4 人、17 歲的有 1 人。計算這個班學生的平均年齡(結果保留到小數點後第一位)。

習　題　十　二

1. 一個中學足球隊的 20 名隊員之身高如下(單位：cm)：

170	167	171	168	160	172	168	162	172	169
164	174	169	165	175	170	165	167	170	172

   計算平均身高(結果保留到個位)。

2. 為了檢查一批零件的質量，從中抽取 10 件，量得它們的長度如下(單位：cm)：

22.36	22.35	22.33	22.35	22.37
22.34	22.38	22.36	22.32	22.35

(1) 在這個問題中，總體、個體、樣本與樣本容量各指什麼？
(2) 計算樣本平均數(結果保留到小數點後第二位)。
(提示：利用公式(2)計算時，可取 $a = 22.30$)。

3. 為了考察一塊實驗林裡樹苗的高度，從中抽取了 20 株，並量得株高如下(單位：cm)：

346	294	365	315	339	313	317	305	321	325
315	329	324	329	368	336	366	308	301	362

(1) 在這個問題中，總體與樣本各指什麼？
(2) 計算樣本平均數(結果保留到個位)。

4. 某農業改良場為了選擇水稻良種，在 10 個試驗點對甲、乙兩個品種作了對比試驗，結果如下：

品種	各試驗點公畝產量(單位：kg)									
	1	2	3	4	5	6	7	8	9	10
甲	779	817	853	793	740	963	794	777	875	864
乙	808	771	725	750	750	859	745	740	706	824

那個品種的平均產量較高？

5. 某班一次語文測驗的成績如下：得 100 分的 7 人、90 分的 14 人、80 分的 17 人、70 分的 8 人、60 分的 2 人、50 分的 2 人。計算這次考試全班的平均成績(結果保留到個位)。

6. 如果兩組數 x_1 、 x_2 、⋯、 x_n 與 y_1 、 y_2 、⋯、 y_n 的平均數分別是 $\overline{x}$ 與 $\overline{y}$ ，那麼一組新數

$$x_1 + y_1 \text{ 、 } x_2 + y_2 \text{ 、 } \cdots \text{ 、 } x_n + y_n$$

的平均數是什麼？為什麼？

16.3 方差

兩台機床同時生產直徑是 40 mm 的零件，從產品中各抽出 10 件進行測量，結果如下(單位：mm)：

機床甲	40	39.8	40.1	40.2	39.9	40	40.2	39.8	40.2	39.8
機床乙	40	40	39.9	40	39.9	40.2	40	40.1	40	39.9

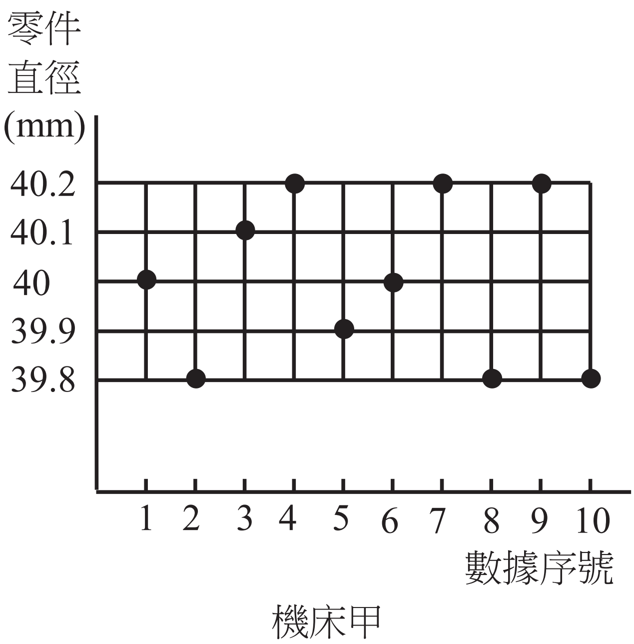
利用公式(2)分別計算這兩組數據的平均數(在公式(2)中取 $a=40$)，得

$$\overline{x_{甲}} = 40 + \frac{1}{10}[0+(-0.2)+\cdots+(-0.2)] = 40$$

$$\overline{x_{乙}} = 40 + \frac{1}{10}[0+0+\cdots+(-0.1)] = 40$$

即這兩組零件直徑數據的平均數都是 40 mm，這時能不能說，在使零件的直徑符合規定方面，兩台機床加工的情況一樣呢？

上面表中的數據如圖 16-1 所示。從圖中看到，機床甲生產零件的直徑與規定尺寸之偏差較大，各點偏離 40 mm 線較多；機床乙生產零件的直徑與規定尺寸之偏差較小，各點比較集中在 40 mm 線的附近。這說明，在使零件的直徑符合規定方面，機床乙比機床甲要好。



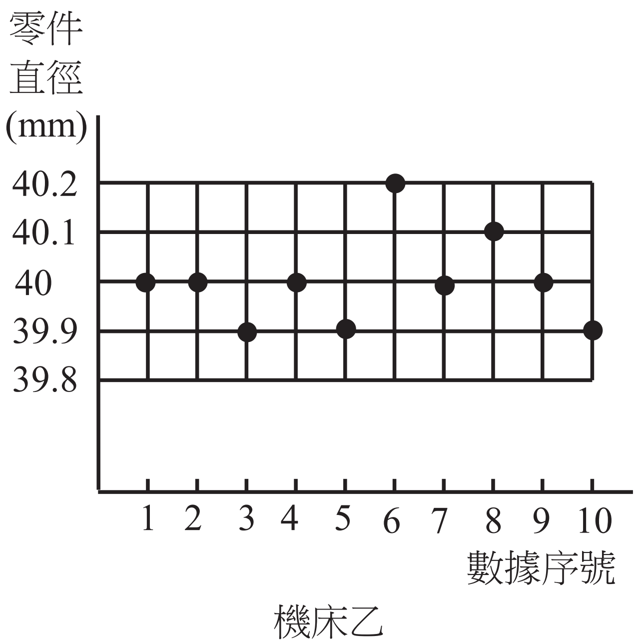


圖 16-1

怎樣用數量來表示樣本的數據偏離平均數之大小呢？在本例中，機床甲生產的 10 個零件直徑的 mm 數分別與樣本平均數 40 的偏差是：

0　–0.2　0.1　0.2　–0.1　0　0.2　–0.2　0.2　–0.2

我們容易想到，是否可以用上面各個偏差的和來衡量零件直徑相對於平均直徑之偏差大小呢？不可以，因為不難驗證，由於正偏差與負偏差相互抵銷，上面各偏差的和等於零。要解決這個矛盾，辦法不只一種，這裡採用將上面各偏差平方後再相加的辦法，這樣，其中各項就不可能是負數了。還有一個問題，各偏差的平方和與樣本容量有關，樣本容量越大，各偏差的平方和也越大。為了排除樣本容量的影響，我們將各偏差平方和再除以樣本容量 n，即用各偏差平方的平均數(用 s^2 表示)來衡量樣本的數據偏離樣本平均數的大小。於是，對於機床甲，

$$\begin{aligned} s^2 &= \frac{1}{10}[(40-40)^2+(39.8-40)^2+\cdots+(39.8-40)^2] \\ &= \frac{1}{10}[0^2+(-0.2)^2+\cdots+(-0.2)^2] \\ &= \frac{1}{10}\times 0.26 \\ &= 0.026 \quad (\mathrm{mm}^2) \end{aligned}$$

對於機床乙，

$$\begin{aligned} s^2 &= \frac{1}{10}[(40-40)^2+(40-40)^2+\cdots+(39.9-40)^2] \\ &= \frac{1}{10}[0^2+0^2+\cdots+(-0.1)^2] \\ &= \frac{1}{10}\times 0.08 \\ &= 0.008 \quad (\mathrm{mm}^2) \end{aligned}$$

由於 $0.026 > 0.008$，說明機床甲生產的零件直徑與規定尺寸偏差較大，機床乙生產的零件直徑與規定尺寸偏差較小。

一般地，樣本中各數據與樣本平均數的差之平方的平均數

$$s^2 = \frac{1}{n}[(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \cdots + (x_n - \overline{x})^2] \qquad (3)$$

叫做**樣本方差**，用來衡量一個樣本的波動大小(即樣本中的數據偏離樣本平均數的大小)。樣本方差越大，說明樣本的波動越大。

為了書寫方便，樣本方差常寫作：

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

當樣本容量很大時，樣本方差很接近反映總體波動大小的特徵數—**總體方差**。通常是用樣本方差去估計總體方差。這樣，通過比較兩個樣本方差，可以近似地比較相應的兩個總體方差之大小。

【例】　已知兩個樣本：

甲：	9.9	10.3	9.8	10.1	10.4	10	9.8	9.7
乙：	10.2	10	9.5	10.3	10.5	9.6	9.8	10.1

分別計算兩個樣本方差。

解　根據公式(2)(取 $a = 10$)，有

$$\begin{aligned}\overline{x_{甲}} &= 10 + \frac{1}{8}(-0.1 + 0.3 - 0.2 + 0.1 + 0.4 + 0 - 0.2 - 0.3)\\ &= 10 + \frac{1}{8}\times 0\\ &= 10\end{aligned}$$

$$\begin{aligned}\overline{x_{乙}} &= 10 + \frac{1}{8}(0.2 + 0 - 0.5 + 0.3 + 0.5 - 0.4 - 0.2 + 0.1)\\ &= 10 + \frac{1}{8}\times 0\\ &= 10\end{aligned}$$

於是，

$$
\begin{aligned}
s_{甲}^2 &= \frac{1}{8}[(9.9-10)^2+(10.3-10)^2+\cdots+(9.7-10)^2] \\
&= \frac{1}{8}[(-0.1)^2+0.3^2+\cdots+(-0.3)^2] \\
&= \frac{1}{8}(0.01+0.09+\cdots+0.09) \\
&= \frac{1}{8}\times 0.44 \\
&= 0.055 \\
s_{乙}^2 &= \frac{1}{8}[(10.2-10)^2+(10-10)^2+\cdots+(10.1-10)^2] \\
&= \frac{1}{8}[0.2^2+0^2+\cdots+0.1^2] \\
&= \frac{1}{8}(0.04+0+\cdots+0.01) \\
&= \frac{1}{8}\times 0.84 \\
&= 0.105
\end{aligned}
$$

從 $s_{甲}^2 < s_{乙}^2$ 知道，樣本乙比樣本甲的波動大。

在有些情況下，需要用到樣本方差的算術平方根

$$\boldsymbol{s=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}} \tag{4}$$

並把它叫做**樣本標準差**，它也是一個用來衡量樣本波動大小的重要之量。樣本數據與樣本方差的度量單位是不一致的。如在本節有關零件直徑的例子中，樣本數據的單位是 mm，而樣本方差的單位是 mm^2。但樣本數據與樣本標準差的度量單位卻是一致的。

在本節有關零件直徑的例子中，兩個樣本標準差分別是：

$$s_{甲}=\sqrt{0.026}\approx 0.16 \quad (mm)$$

$$s_{乙}=\sqrt{0.008}\approx 0.089 \quad (mm)$$

練　習

1. 計算下列各樣本方差與標準差(結果保留到小數點後第一位)

   (1)　−2、2、0、−3、−2、3、0、1；

   (2)　28、24、25、23、27、24、22、24、25、28。

2. 從甲、乙兩名車工車出的同一種零件中，各抽出 5 個，量得它們的直徑(單位：mm)如下(圖紙上規定的尺寸是 10 mm)：

甲生產的零件尺寸	10.05	10.02	9.97	9.96	10.00
乙生產的零件尺寸	10.00	10.01	10.02	9.97	10.00

   分別計算上面兩個樣本的平均數與方差，說明在使零件的直徑符合規定方面，誰做得較好。

16.4　方差的簡化計算[8]

按照公式(3)計算樣本方差，計算量比較大，下面介紹方差的簡化計算方法。

為了簡化討論，我們假定樣本中僅包含 3 個數據 x_1、x_2、x_3，它們的平均數是 $\bar{x}$ ，那麼樣本方差是

$$s^2=\frac{1}{3}[(x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2]\text{ 。}$$

將方括號內的各項展開後再整理，得到

$$\begin{aligned}
s^2&=\frac{1}{3}[(x_1^2-2x_1\bar{x}+\bar{x}^2)+(x_2^2-2x_2\bar{x}+\bar{x}^2)+(x_3^2-2x_3\bar{x}+\bar{x}^2)]\\
&=\frac{1}{3}[(x_1^2+x_2^2+x_3^2)-2(x_1+x_2+x_3)\bar{x}+3\bar{x}^2]\\
&=\frac{1}{3}[(x_1^2+x_2^2+x_3^2)-2\times3\times(\frac{x_1+x_2+x_3}{3})\bar{x}+3\bar{x}^2]\\
&=\frac{1}{3}[(x_1^2+x_2^2+x_3^2)-2\times3\bar{x}^2+3\bar{x}^2]\\
&=\frac{1}{3}[(x_1^2+x_2^2+x_3^2)-3\bar{x}^2]
\end{aligned}$$

[8] 方差的簡化計算公式不作為要求。

一般地，如果樣本的容量是 n，那麼樣本方差可以用下面的公式計算：

$$\boxed{s^2=\frac{1}{n}[(x_1^2+x_2^2+\cdots+x_n^2)-n\bar{x}^2]} \qquad (5)$$

或者簡記作

$$\boxed{s^2=\frac{1}{n}\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right)} \qquad (5)$$

用公式(5)計算樣本方差，是直接計算各個數據的平方，而不必計算各個數據與樣本平均數的差之平方，因此它比用公式(3)計算少一個步驟，有時比較方便。

【例 1】 計算下面樣本的方差(結果保留到小數點後第一位)：

3、 –1 、2、1、 –3 、3

解　從計算知道，樣本平均數不是整數。這時用公式(3)計算樣本方差比較麻煩，我們用公式(5)來計算。

$$\begin{aligned}
s^2&=\frac{1}{6}\left[3^2+(-1)^2+2^2+1^2+(-3)^2+3^2-6\times\left(\frac{3-1+2+1-3+3}{6}\right)^2\right]\\
&=\frac{1}{6}\left[9+1+4+1+9+9-6\times\left(\frac{5}{6}\right)^2\right]\\
&=\frac{1}{6}\left(33-6\times\frac{25}{36}\right)\\
&=\frac{1}{6}\times 33-\frac{25}{36}\\
&\approx 5.5-0.7\\
&=4.8
\end{aligned}$$

練　習

計算下面樣本的方差(結果保留到小數點後第一位)：

5、4、4、3、4、3、2、3、5、3。

當樣本數據較大時，用公式(5)計算樣本方差仍然比較麻煩。如果數據相互比較接近，為了減小參與數據的計算，我們可以仿照前面簡化計算平均數的方法，將每一個數據減去一個與樣本平均數接近的常數 a。假定樣本中包含 3 個數據 x_1 、 x_2 、 x_3 ，我們作替換

$$x_1' = x_1 - a \text{ 、 } x_2' = x_2 - a \text{ 、 } x_3' = x_3 - a \text{ ，}$$

那麼

$$x_1 = x_1' + a \text{ 、 } x_2 = x_2' + a \text{ 、 } x_3 = x_3' + a \text{ 。}$$

又根據公式(2)，有

$$\bar{x} = \bar{x}' + a \text{ ，}$$

其中 $\bar{x}'$ 是 x_1' 、 x_2' 、 x_3' 的平均數，於是，

$$\begin{aligned} s^2 &= \frac{1}{3}[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_2 - \bar{x})^2] \\ &= \frac{1}{3}\{[(x_1' + a) - (\bar{x}' + a)]^2 + [(x_2' + a) - (\bar{x}' + a)]^2 + [(x_3' + a) - (\bar{x}' + a)]^2\} \\ &= \frac{1}{3}[(x_1' - \bar{x}')^2 + (x_2' - \bar{x}')^2 + (x_3' - \bar{x}')^2] \\ &= \frac{1}{3}[(x_1'^2 + x_2'^2 + x_3'^2) - 3\bar{x}'^2] \quad \text{（根據公式(5)）} \end{aligned}$$

一般地，如果樣本的容量是 n，那麼有

$$\boxed{\boldsymbol{s^2 = \frac{1}{n}[(x_1'^2 + x_2'^2 + \cdots + x_n'^2) - n\bar{x}'^2]}} \qquad (6)$$

其中

$$x_1' = x_1 - a \text{ 、 } x_2' = x_2 - a \text{ 、} \cdots \text{ 、 } x_n' = x_n - a \text{ ，}$$

這裡 a 是接近樣本平均數的常數。

上式可以簡記為

$$\boxed{\boldsymbol{s^2 = \frac{1}{n}\left(\sum_{i=1}^{n} x_i'^2 - n\bar{x}'^2\right)}} \qquad (6')$$

由於在替換$x_i'=x_i-a$ ($i=1$、2、⋯、n)中所取的常數 a 與樣本平均數$\bar{x}$比較接近，各數據x_i'比較小，因此當樣本數據大時，用公式(6)計算樣本方差比較簡便。

【例 2】從小麥田裡抽取了甲、乙兩種小麥各 10 株，測得各株高如下(單位：cm)：

甲：	76	90	84	86	81	87	86	82	85	83
乙：	82	84	85	89	79	80	91	89	79	74

哪種小麥長的比較整齊？

解　根據題意，是要比較兩個樣本方差的大小。因為樣本數據較大，我們用公式(6)來計算兩個樣本方差，步驟如下：

(1) 確定替換$x_i'=x_i-a$中的常數 a。

由於樣本數據都在 80 左右波動，我們取$a=80$ 。

(2) 如表 1、表 2 所示，分別計算各個x_i'及$x_i'^2$ ($i=1$、2、⋯、10)，然後計算各個x_i'的和及各個$x_i'^2$的和，並填入表中。

表 1　(甲種小麥)

x_i	x_i' (x_i-80)	$x_i'^2$
76	–4	16
90	10	100
84	4	16
86	6	36
81	1	1
87	7	49
86	6	36
82	2	4
85	5	25
83	3	9
合計	40	292

表 2　(乙種小麥)

x_i	x_i' (x_i-80)	$x_i'^2$
82	2	4
84	4	16
85	5	25
89	9	81
79	–1	1
80	0	0
91	11	121
89	9	81
79	–1	1
74	–6	36
合計	32	366

(3) 將表中有關數據代入公式(6)進行計算：

$$
\begin{aligned}
s_{甲}^2 &= \frac{1}{10}\left(\sum_{i=1}^{10} x_i'^2 - 10\overline{x}'^2\right)\\
&= \frac{1}{10}\left[292 - 10\times\left(\frac{40}{10}\right)^2\right]\\
&= \frac{1}{10}(292-160)\\
&= \frac{1}{10}\times 132\\
&= 13.2
\end{aligned}
$$

$$
\begin{aligned}
s_{乙}^2 &= \frac{1}{10}\left(\sum_{i=1}^{10} x_i'^2 - 10\overline{x}'^2\right)\\
&= \frac{1}{10}\left[366 - 10\times\left(\frac{32}{10}\right)^2\right]\\
&= \frac{1}{10}(366-102.4)\\
&= \frac{1}{10}\times 263.6\\
&= 26.36
\end{aligned}
$$

因為 $s_{甲}^2 < s_{乙}^2$ ，可以估計，甲種小麥比乙種小麥長的整齊。

【例 3】 某農試所在 8 個試驗點對某兩個水稻品種進行栽培比對試驗，兩個品種在各試驗點的畝產量如下(單位：kg)：

甲：	804	984	989	817	919	840	912	1001
乙：	856	932	930	855	872	910	897	918

哪個品種的產量比較穩定？

解 根據題意，是要比較兩個樣本方差的大小。由於樣本數據較大，我們仿照例 2 的步驟，根據公式(6)來計算兩個樣本方差。

(1) 由於兩個樣本中的數據都在 900 左右波動，我們在替換 $x_i' = x_i - a$ 中，取 $a = 900$ 。

(2) 如表 3、表 4 所示，分別計算各個 x_i' 及 $x_i'^2$ ，然後計算各個 x_i' 的和及各個 $x_i'^2$ 的和，並填入表中。

表 3　(甲種水稻)

x_i	x_i' $(x_i - 900)$	$x_i'^2$
804	–96	9216
984	84	7056
989	89	7921
817	–83	6889
919	19	361
840	–60	3600
912	12	144
1001	101	10201
合計	66	45388

表 4　(乙種水稻)

x_i	x_i' $(x_i - 900)$	$x_i'^2$
856	–44	1936
921	32	1024
930	30	900
855	–45	2025
872	–28	784
910	10	100
897	–3	9
918	18	324
合計	–30	7102

(3) 將表中有關數據代入公式(6)進行計算：

$$s_{甲}^2 = \frac{1}{8}\left(\sum_{i=1}^{8} x_i'^2 - 8\overline{x}'^2\right)$$

$$= \frac{1}{8}\left[45388 - 8\times\left(\frac{66}{8}\right)^2\right]$$

$$= \frac{1}{8}(45388 - 544.5)$$

$$= \frac{1}{8}\times 44843.5$$

$$\approx 5605$$

$$s_{乙}^2 = \frac{1}{8}\left(\sum_{i=1}^{8} x_i'^2 - 8\overline{x}'^2\right)$$

$$=\frac{1}{8}\left[7102-8\times\left(\frac{-30}{8}\right)^2\right]$$

$$=\frac{1}{8}(7102-112.5)$$

$$=\frac{1}{8}\times 6989.5$$

$$\approx 874$$

因為$s_{乙}^2 < s_{甲}^2$，可以估計，乙種水稻比甲種水稻的產量穩定。

我們看到，樣本平均數與方差的計算往往較繁。為便於計算，所以樣本容量應盡可能「整」一點，如 10、20、30 等。

練 習

計算下面各樣本方差(結果保留到小數點後第一位)：

(1) 105、103、101、100、114、108、110、106、98、102；

(2) 423、421、419、420、421、417、422、419、423、418。

習 題 十 三

1. 甲、乙兩人在相同條件下各射靶 10 次，各次命中的環數如下：

甲：	7	8	6	8	6	5	9	10	7	4
乙：	9	5	7	8	7	6	8	6	7	7

分別計算上面兩個樣本方差。從計算結果看，誰射擊的情況比較穩定？

2. 甲、乙兩台機器同時生產一種零件。在 10 天中，兩台機器每天出的瑕疵品數分別是：

甲：	0	1	0	2	2	0	3	1	2	4
乙：	2	3	1	1	0	2	1	1	0	1

分別計算上面兩個樣本的平均數與方差。從計算結果看，哪台機器的性能較好？

3. 某農場種植的甲、乙兩種水稻，在連續六年中各年的平均畝產量如下(單位：kg)：

品種	第 1 年	第 2 年	第 3 年	第 4 年	第 5 年	第 6 年
甲：	900	920	900	850	910	920
乙：	890	960	950	850	860	890

哪種水稻產量比較穩定？

4. 下面是今年與前年在大致相同條件下飼養的十頭豬之體長數據(單位：cm)：

前年： 112 110 110 117 113 122 125 124 119 127
今年： 111 122 115 123 114 115 118 114 116 115

哪年飼養的豬之體長比較一致？

16.5 頻率分布

為了了解中學生的身體發育情況，對某中學同年齡的 60 名女學生之身高進行了測量，結果如下(單位：cm)：

167	154	159	166	169	159	156	166	162	158
159	156	166	160	164	160	157	156	157	161
158	158	153	158	164	158	163	158	153	157
162	162	159	154	165	166	157	151	146	151
158	160	165	158	163	163	162	161	154	165
162	162	159	157	159	149	164	168	159	153

我們知道，這組數據的平均數，反映了這些學生的平均身高，但是，有時只知道這一點還不夠，還希望知道身高在哪個小範圍內的學生多，在哪個小範圍的學生少，也就是希望知道這 60 名女學生的身高數據在各個小範圍所佔的比例大小。為此，需要對這組數據進行適當整理。整理時，可以按照下面的步驟進行：

(1)　計算最大值與最小值的差。

在上面的數據中，最大值是 169、最小值是 146，它們的差是

$$169-146=23\,(\text{cm})。$$

算出了最大值與最小值的差，就知道這組數據變動的範圍有多大。

(2)　決定組距與組數。

將一批數據分組，一般數據越多，分的組數也越多。當數據在 100 個以內時，按照數據的多少，常分成 5～12 組。

組距是指每個小組的兩個端點之間的距離。在本例中，如果取組距為 3 cm，那麼由於在這批數據中，

$$\frac{\text{最大值}-\text{最小值}}{\text{組距}}=\frac{23}{3}=7\frac{2}{3}\;,$$

得將數據分成 8 組；如果取組距為 2 cm，那麼由於 $\frac{23}{2}=11\frac{1}{2}$，得分成 12 組。若分成 8 組較合適，可取組距為 3 cm、組數為 8。

(3)　決定分點。

將數據按照 3 cm 的組距分組時，可以分成以下 8 組：

146～149、149～152、152～155、155～158、
158～161、161～164、164～167、167～170。

這時我們看到，有些數據(例如 149、158、167)本身就是分點，不好決定它們究竟應該屬於哪一組。為了避免出現這種情況，可以使分點比數據多一位小數，並且把第 1 組的起點稍微減小一點。例如，可以將第 1 組的起點訂為 145.5，這樣，所分的 8 個小組是：

145.5～148.5、148.5～151.5、151.5～154.5、154.5～157.5、
157.5～160.5、160.5～163.5、163.5～166.5、166.5～169.5。

(4)　列頻率分布表。

如表 5 的第 1 行、第 2 行所示，用選舉時唱票的方法，對落在各個小組內的數據進行累計。然後，數出落在各個小組內的數據之個數(叫做**頻數**)，並且填入表 5 中的第 3 行。

表 5　頻率分布表

分　組	頻 數 累 計	頻　數	頻　率
145.5～148.5	一	1	0.017
148.5～151.5	下	3	0.050
151.5～154.5	正一	6	0.100
154.5～157.5	正下	8	0.133
157.5～160.5	正正正下	18	0.300
160.5～163.5	正正一	11	0.183
163.5～166.5	正正	10	0.167
166.5～169.5	下	3	0.050
合計		60	1.000

每一小組的頻數與樣本容量之比值叫做這一小組的**頻率**。例如，第 1 小組的頻率是

$$\frac{1}{60} \approx 0.017\text{ 。}$$

算出各個小組的頻率，並填入表 5 的第 4 行。表 5 叫做**頻率分布表**。列出頻率分布表以後，就知道這些數據在各個小組內所佔的比例大小了。

(5)　繪製頻率分布直方圖。

為了將頻率分布表中的結果直觀形象地表示出來，常繪出**頻率分布直方圖**。對於本例，頻率分布直方圖如圖 16-2 所示，其中橫軸表示身高、縱軸表示頻率與組距的比值。容易看出，

$$\text{小長方形面積} = \text{組距} \times \frac{\text{頻率}}{\text{組距}} = \text{頻率}\text{ ，}$$

這就是說，各個小長方形的面積等於相應各組之頻率。這樣，頻率分布直方圖就以圖形面積的形式反映了數據落在各個小組內的頻率之大小。又在圖 16-2 中，

$$\text{小長方形高} = \frac{\text{頻率}}{\text{組距}} = \frac{1}{\text{組距} \times \text{樣本容量}} \times \text{頻數}\text{ ，}$$

因為組距與樣本容量都是常數，$\dfrac{1}{\text{組距}\times\text{樣本容量}}$ 也是常數，所以小長方形的高與頻數成正比。利用這個性質來確定各小長方形的高，那麼頻數為 k 的小長方形之高就是 kh。例如，148.5～151.5 這個小組的頻數是 3，相應的小長方形之高就是 $3h$。

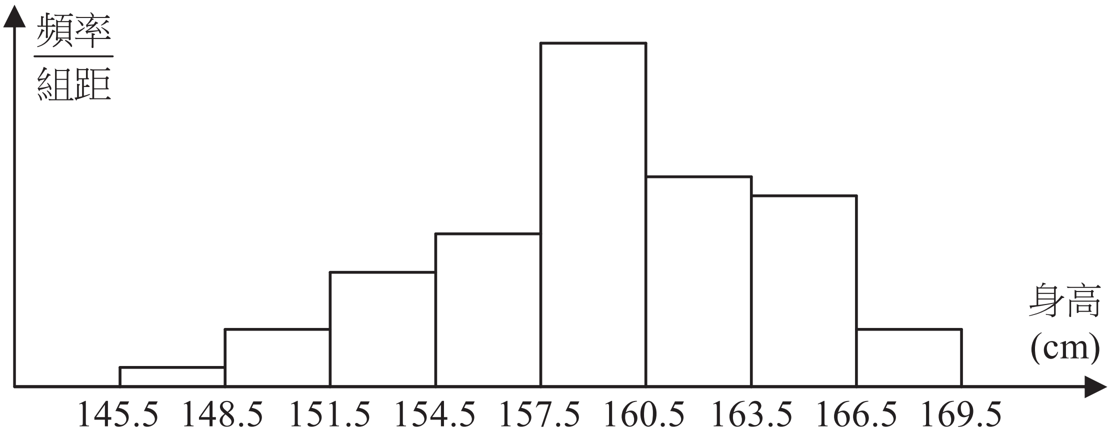


圖 16-2

在頻率分布直方圖中，由各小長方形的面積等於相應各組之頻率，而各組的頻率之和等於 1，故各小長方形的面積之和為 1。

知道了一個樣本的頻率分布以後，就可以對相應的總體分布做出估計。在上面的例子中，樣本數據落在 157.5～160.5 cm 之間的頻率 0.3，說明在每 100 名該年齡的女學生中，約有 30 人的身高在 157.5～160.5 cm 之間。

【例】 一個農試所為了考察某種大麥穗長的分布情況，在一塊試驗田裡抽取了 100 個穗，量得長度如下(單位：cm)：

6.5	6.4	6.7	5.8	5.9	5.9	5.2	4.0	5.4	4.6
5.8	5.5	6.0	6.5	5.1	6.5	5.3	5.9	5.5	5.8
6.2	5.4	5.0	5.0	6.8	6.0	5.0	5.7	6.0	5.5
6.8	6.0	6.3	5.5	5.0	6.3	5.2	6.0	7.0	6.4
6.4	5.8	5.9	5.7	6.8	6.6	6.0	6.4	5.7	7.4
6.0	5.4	6.5	6.0	6.8	5.8	6.3	6.0	6.3	5.6
5.3	6.4	5.7	6.7	6.2	5.6	6.0	6.7	6.7	6.0
5.5	6.2	6.1	5.3	6.2	6.8	6.6	4.7	5.7	5.7
5.8	5.3	7.0	6.0	6.0	5.9	5.4	6.0	5.2	6.0
6.3	5.7	6.8	6.1	4.35	5.6	6.3	6.0	5.8	6.3

列出樣本的頻率分布表，繪出頻率分布直方圖。

解　(1) 計算最大值與最小值的差。

在樣本數據中，最大值是 7.4、最小值是 4.0，差是

$$7.4-4.0=3.4\ (\text{cm})。$$

(2) 決定組距與組數。

在本例中，最大值與最小值的差是 3.4 cm。如果取組距為 0.3 cm，那麼由於 $\frac{3.4}{0.3}=11\frac{1}{3}$，得分成 12 組，組數合適。於是選定組距為 0.3 cm，組數為 12。

(3) 決定分點。

使分點比數據多一位小數，並且把第 1 小組的起點稍微減小一點，那麼，所分的 12 小組可以是：

3.95～4.25、4.25～4.55、4.55～4.85、4.85～5.15、5.15～5.45、5.45～5.75、5.75～6.05、6.05～6.35、6.35～6.65、6.65～6.95、6.95～7.25、7.25～7.55。

(4) 列頻率分布表。

表 6　頻率分布表

分　組	頻 數 累 計	頻　數	頻　率	累積頻率[9]
3.95～4.25	一	1	0.01	0.01
4.25～4.55	一	1	0.01	0.02
4.55～4.85	丅	2	0.02	0.04
4.85～5.15	正	5	0.05	0.09
5.15～5.45	正正一	11	0.11	0.20
5.45～5.75	正正正	15	0.15	0.35
5.75～6.05	正正正正正下	28	0.28	0.63
6.05～6.35	正正下	13	0.13	0.76
6.35～6.65	正正一	11	0.11	0.87
6.65～6.95	正正	10	0.10	0.97
6.95～7.25	丅	2	0.02	0.99
7.25～7.55	一	1	0.01	1.00
合　計		100	1.000	

[9] 關於累積頻率與圖 16-3 的下圖將在下一節作專門介紹

對各個小組作頻數累計，然後數頻數、算頻率、列頻率分析表，如表 6 所示。

(5) 繪製頻率分布直方圖(如圖 16-3 上圖所示，繪圖方法同上例)。

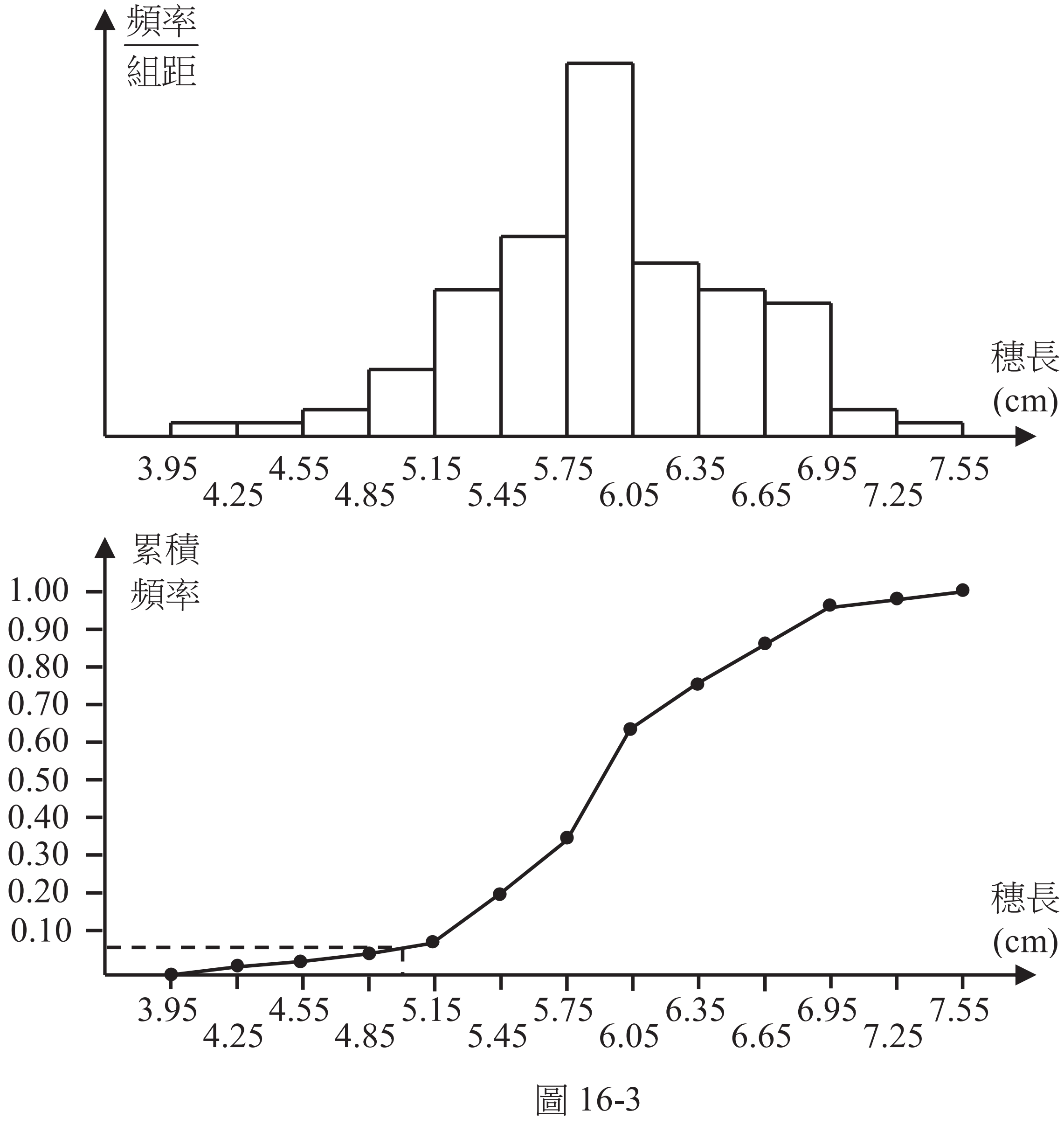


圖 16-3

現在可以根據樣本的頻率分布來估計總體分布。例如從表 6 中看到，樣本數據落在 5.75～6.05 之間的頻率是 0.28，於是可以估計，在這塊田裡，長度在 5.75～6.05 cm 之間的麥穗約佔 28%。在用頻率分布估計總體分布時，如果樣本容量越大，這種估計也就越精確。

練　習

1. 有一個容量為 50 的樣本，數據的分組以及各組之頻數如下：

53.5～55.5	4	61.5～63.5	10
55.5～57.5	7	63.5～65.5	6
57.5～59.5	9	65.5～67.5	3
59.5～61.5	11		

列出樣本的頻率分析表，繪出頻率分布直方圖。

2. 已知一個樣本：

25	21	23	25	27	29	25	28	30	29
26	24	25	27	26	22	24	25	26	28

填寫下面的頻率分布表：

分　組	頻數累計	頻　數	頻　率
20.5～22.5			
22.5～24.5			
24.5～26.5			
26.5～28.5			
28.5～30.5			
合　計			

3. 為了了解中學生的身體發育情況，對某一中學同年齡的 50 名男同學的身高進行了測量，結果如下(單位：cm)

175	168	170	176	167	181	162	173	171	177
179	172	165	157	172	173	166	177	169	181
160	163	166	177	175	174	173	174	171	171
158	170	165	175	165	174	169	163	166	166
174	172	166	172	167	172	175	161	173	167

列出樣本的頻率分析表，繪出頻率分布直方圖。

*16.6 累積頻率分布[10]

在上節關於大麥穗長的例題中，穗長小於 4.55 cm 的頻率等於前兩個小組的頻率之和，即

$$0.01+0.01=0.02 \text{。}$$

穗長小於 4.85 cm 的頻率等於前三個小組的頻率之和，即

$$0.01+0.01+0.02=0.04 \text{。}$$

依此類推。這種數據小於某一數值的頻率叫做該數值的**累積頻率**。這個例子中各個分點的累積頻率如表 6 最後一行所示。它可以作為頻率分布表的一個補充。

根據算出累積頻率，可以繪出**累積頻率分布圖**，如圖 16-3 下圖所示。圖中的橫軸表示穗長、縱軸表示累積頻率。按照表中的各累積頻率，在圖中描出相應的各點。例如，分點 4.25 的累積頻率是 0.01，就在圖中描出橫座標是 4.25、縱座標是 0.01 的一個點。然後，用線段將各點依次連結起來，就得到一條折線，它就是累積頻率分布圖

利用樣本的累積頻率分布圖，可以對總體的相應情況做出估計。例如，為了估計長度小於 5 cm 的麥穗在這塊麥地裡所佔之比例大小，我們在累積頻率分布圖找到橫座標是 5 (cm)的點，然後量出這一點的縱座標 0.07。它告訴我們，長度小於 5 cm 的麥穗在這塊麥地裡約佔 7%。

練 習

根據表 5，計算各分點的累積頻率，並繪出累積頻率分布圖。

[10] 本節是選學內容。

習 題 十 四

1. 在一批木竿中抽測了 60 根木竿長度，結果如下(單位：mm)：

82	202	352	321	25	293	293	86	28	206
323	355	357	33	325	113	233	294	50	296
115	236	357	326	52	301	140	328	238	358
58	255	143	360	340	302	370	343	260	303
59	146	60	263	170	305	380	346	61	305
175	348	264	383	62	306	195	350	265	385

列出樣本的頻率分析表，繪出頻率分布直方圖。

2. 從某批機器零件中抽測了 100 個，其尺寸與規定尺寸的偏差如下(單位：0.1 mm)：

2	1	0	3	–1	2	1	0	–1	0
3	–2	–1	–2	1	0	0	–4	1	0
1	2	2	0	–2	1	0	–1	0	3
–1	1	0	–1	0	3	1	–2	3	–1
–3	0	0	1	4	0	–2	2	1	2
2	0	–2	0	0	–1	1	4	–2	1
1	1	4	–1	1	–1	0	2	–2	1
0	1	–1	1	0	2	2	1	0	–1
0	–1	3	1	2	–3	1	0	1	1
–2	2	1	2	–1	–2	3	1	–1	0

(1) 列出樣本的頻率分析表，繪出頻率分布直方圖(提示：由於樣本中只有 9 個不同的數值，且它們是連續的整數，所以最多只能將數據分成 9 個小組)；

(2) 根據樣本的頻率分布表，估計整批機器中與規定尺寸的偏差之絕對值不超過 0.1 mm 的約佔多少，不超過 0.2 mm 的約佔多少。

3. 一個農場在秋收前進行玉米估產，從地裡抽取了 100 株玉米，稱得各株玉米的產量如下(單位：g)：

50	29	31	32	29	30	38	41	26	20
33	27	34	20	27	17	17	28	35	39
38	35	26	50	38	36	38	25	28	22
32	30	29	48	30	35	18	20	25	24
29	50	23	44	39	27	33	39	24	26
34	23	32	18	39	25	50	33	40	21
35	31	33	30	28	25	27	30	39	24
35	30	38	46	40	18	35	43	24	23
33	34	36	34	40	23	34	37	40	19
39	34	33	37	35	25	34	30	27	15

(1) 列出樣本的頻率分析表，繪出頻率分布直方圖；

*(2) 在頻率分布表中補填各分點的累積頻率，並繪出累積頻率分布圖。

小　结

一、這一章我們介紹了統計的一些初步知識。統計方法的特點是：從所要考察的總體中抽取一個樣本，通過研究樣本對總體做出估計。樣本容量越大，這種估計也就越精確。。

二、總體平均數式表示總體的平均大小(或平均水平)之特徵數，通常用樣本平均數去估計它。樣本平均數是指樣本中各數據的和與樣本容量之比值，當樣本數據較大時，利用公式 $\bar{x}=\bar{x}'+a$ 計算樣本平均數比較簡便。

三、總體方差式表示總體的波動大小之特徵數，通常用樣本方差去估計它，用比較兩個樣本方差去近似地比較兩個相應的總體之波動大小。對於一個容量為 n、平均數為 $\bar{x}$ 的樣本來說，樣本方差

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \text{ 。}$$

樣本方差的計算較繁，在沒有電子計算器的情況下，一般根據簡化公式列表進行計算。

四、總體分布反映了總體中各部分個體在總體中所佔的比例大小，通常用樣本的頻率分布去估計它。頻率分布反映了樣本數據在各個小範圍內所佔的比例大小，要得到一個樣本的頻率分布情況，可按下列步驟進行：計算樣本中最大值與最小值的差、決定組距與組數、決定分點、列頻率分布表、繪頻率分布直方圖。

複習參考題十六

1. 什麼叫做總體？什麼叫做總體的一個樣本？為什麼常常需要用樣本的某些特性去估計總體的相應特性？

2. 甲、乙兩門砲在相同條件下向同一目標各發射 50 發砲彈，砲彈落點情況如下表所示：

砲彈落點與目標距離	40 m	30 m	20 m	10 m	0 m
甲砲發射的砲彈個數	0	1	3	7	39
乙砲發射的砲彈個數	1	3	2	3	41

分別計算兩個樣本平均數，根據計算結果，估計哪門砲射擊的準確性較好。

3. 一個水庫養了某種魚 10 萬條，從中撈捕了 20 條，稱得它們的體重如下(單位：kg)

2.3	2.1	2.2	2.1	2.2	2.6	2.5	2.4	2.3	2.4
2.4	2.3	2.2	2.5	2.4	2.6	2.3	2.5	2.2	2.3

計算樣本平均數，並根據計算結果估計水庫裡約有這種魚多少 kg。

4. 為了考察甲、乙兩種小麥的生長情況，分別從中抽取 10 株苗，測得苗高如下(單位：cm)：

甲：	12	13	14	15	10	16	13	11	15	11
乙：	11	16	17	14	13	19	6	8	10	16

(1) 分別計算兩種小麥的平均苗高；

(2) 哪種小麥長的比較整齊？

5. 兩名跳遠運動員在 10 次比賽中的成績分別如下(單位：m)：

甲：	5.85	5.93	6.07	5.91	5.99	6.13	5.98	6.05	6.00	6.19
乙：	6.11	6.08	5.83	5.92	5.84	5.81	6.18	6.17	5.85	6.21

分別計算兩個樣本方差，並根據計算結果估計哪名運動員的成績比較穩定。

6. 從一種零件中抽取了 80 件，尺寸數據如下(單位：m)：

362.5×1	362.6×2	362.7×2	362.8×3
362.9×3	363.0×3	363.1×5	363.2×6
363.3×8	363.4×9	363.5×9	363.6×7
363.7×6	363.8×4	363.9×3	364.0×3
364.1×2	364.2×2	364.3×1	364.4×1

(1) 列出樣本的頻率分析表，繪出頻率分布直方圖；

*(2)在頻率分布表中加填各分點的累積頻率，並繪出累積頻率分布圖。

7. (1) 設 $\overline{x}$ 是 x_1、x_2、⋯、x_n 的平均數、$\overline{y}$ 是 $3x_1+5$、$3x_2+5$、⋯、$3x_n+5$ 的平均數，求證 $\overline{y}=3\overline{x}+5$ 。

(2) 設 s_x 是 x_1、x_2、⋯、x_n 的標準差、s_y 是 $3x_1+5$、$3x_2+5$、⋯、$3x_n+5$ 的標準差，求證 $s_y=3s_x$ 。

附　錄　初中代數總複習參考題[11]

1. 列出實數的分類表。
2. 計算：
   (1) $4\times 0.2^2-10\times 1.1^3+3\times\sqrt{5.14}-\sqrt[3]{5.26}$ (保留三個有效數字)；
   (2) $3a^3-[-5a^2+a(-4a^2+2s-1)+4]-7$ (精確到 0.01)，其中 $a=2.35$。
3. (1) 若 $(a-1)^2+(b+2)^2=0$，a、b 為實數，求 a、b；
   (2) 若 $x^2+2x+y^2-6y+10=0$，x、y 為實數，求 x、y。
4. 計算下列各題：
   (1) $(a^3-a^2b+ab^2-b^3)(a+b)$；
   (2) $(a+b)^2(a-b)^2-(a^2+b^2)(a^2-b^2)$；
   (3) $(p+q-m-n)(p-q-m+n)$；
   (4) $(x+2y-z)(x-2y+z)-(x+2y+z)^2$；
   (5) $(x^5+x^4+x^3+x^2+x+1)\div(x+1)$；
   (6) $(x^4+x^2+1)\div(x^2-x+1)$。
5. (1) 已知 $x^2+x-1=0$，求 x^3+2x^2+3 的值；
   (2) 若 $bc=ad$，求證：$ab(c^2-d^2)=(a^2-b^2)cd$；
   (3) 若 $a+b+c=0$，求證：$a^3+a^2c+b^2c-abc+b^3=0$；
   (4) 證明：
   $$a^2(b-c)+b^2(c-a)+c^2(a-b)=-(a-b)(b-c)(c-a)；$$
6. 分解因式：
   (1) $x^5-x^4+x^3-x^2+x-1$；　(2) $3x^2-2x-8$；
   (3) $4x^2+2x+\frac{1}{4}-y^2$；　(4) l^4-3l^2+1；
   (5) $a^2-2ab+b^2-6a+6b+5$；　(6) $(1-a^2)(1-b^2)-4ab$。

---

[11] 這是供複習初中代數時參考選用的習題

7. 在實數集合內分解因式：

(1) x^2+x-1；　(2) x^4+x^2-6；

(3) $6x^4-7x^2-3$；　(4) $x^4+3x^3+x^2$。

8. x 取什麼值時，代數式 $\dfrac{-(x-5)}{(3-x)(x+1)}$ 的值滿足下列條件？

(1) 等於 0；　(2) 沒意義；　(3) 等於 1。

9. 計算下列各題：

(1) $\dfrac{x^2+2x+1}{x^3-x}\times\dfrac{x}{x+1}-\dfrac{1}{x+1}$；

(2) $\dfrac{1}{1-x}+\dfrac{1}{1+x}+\dfrac{2}{1+x^2}+\dfrac{4}{1+x^4}$；

(3) $\dfrac{x+2}{x+1}-\dfrac{x+3}{x+2}+\dfrac{x+4}{x+3}-\dfrac{x+5}{x+4}$；

(4) $\dfrac{a^2}{(a-b)(a-c)}+\dfrac{b^2}{(b-a)(b-c)}+\dfrac{c^2}{(c-a)(c-b)}$。

10. 試證下列各題：

(1) 若 $\dfrac{a}{b}+\dfrac{b}{a}=x$、$\dfrac{a}{b}-\dfrac{b}{a}=y$，求證：$x^2-y^2=4$；

(2) 若 a、b、c 為實數，且 $a^2+b^2+c^2-ab-bc-ca=0$，求證：$a=b=c$。

11. 計算：

(1) $5\sqrt{28}+3\sqrt{63}-10\sqrt{7}+3\sqrt{\dfrac{1}{7}}$；

(2) $4\sqrt[6]{125}-3\sqrt[4]{25}+\sqrt{\dfrac{4}{5}}$；

(3) $(2\sqrt{3}+\sqrt{6})(\sqrt{3}-2\sqrt{6})$；

(4) $\sqrt{(\sqrt{2}-\sqrt{3})^2}+\sqrt{(\sqrt{3}-\sqrt{2})^2}$；

(5) $\sqrt{xy}\left(\sqrt{xy}-3\sqrt{\dfrac{y}{x}}-2\sqrt{\dfrac{x}{y}}+4\sqrt{\dfrac{1}{xy}}\right)$ ($x>0$ 、 $y>0$)；

(6) $\dfrac{1}{1-\sqrt[4]{x}}+\dfrac{1}{1+\sqrt[4]{x}}+\dfrac{2}{1+\sqrt{x}}+\dfrac{4}{1+x}$ 。

12. 化簡 $\sqrt{\dfrac{x^2-6x+9}{x^2+6x+9}}$ 。

13. 解下列方程：

(1) $\dfrac{2-x}{6}-\dfrac{2x-3}{4}=1$ ；　　(2) $\dfrac{3x+6}{8}-\left(\dfrac{5x}{6}-1\right)=\dfrac{5}{6}$ ；

(3) $\dfrac{1}{3}\left(3x-\dfrac{10-7x}{2}\right)-\dfrac{1}{6}\left(2x-\dfrac{2x+2}{3}\right)=\dfrac{x}{2}-1$ ；

(4) $\dfrac{1}{3}(x-2)-\dfrac{1}{7}(5x-6)=\dfrac{22x-63}{105}-\dfrac{1}{5}(3x-4)$ ；

(5) $\dfrac{ax}{b}+\dfrac{bx}{a}-1=0$ (a、b 為已知數) ；

(6) $\dfrac{mx+1}{n}+\dfrac{nx+1}{m}=1$ (m、n 為已知數)。

14. 解下列方程組：

(1) $\begin{cases}\dfrac{1}{2}(x+11)=\dfrac{1}{3}(y+13)+2\\ 5x=3y+8\end{cases}$ ；

(2) $\begin{cases}\dfrac{2x}{3}+\dfrac{3y}{4}=\dfrac{3x-2y}{2}+1\\ \dfrac{3x}{2}-\dfrac{4y}{3}=\dfrac{3x+4y}{6}-1\end{cases}$ ；　　(3) $\begin{cases}\dfrac{3}{x-4}+\dfrac{4}{y-1}=3\\ \dfrac{9}{x-4}-\dfrac{2}{y-1}=2\end{cases}$ ；

(4) $\dfrac{2x+y+6}{4}=\dfrac{4x-3y-7}{6}=\dfrac{-6x-7y+10}{8}$ ；

(5) $\begin{cases} 2x+3y-4z=4 \\ 2y+3z=\dfrac{17}{12} \\ x+4y=\dfrac{10}{3} \end{cases}$；　　(6) $\begin{cases} x+y-z=3 \\ z+x-y=1 \\ y+z-x=7 \end{cases}$。

15. 解方程：

(1) $4(x+3)^2-9(x-2)^2=0$；

(2) $\sqrt{3}(y^2-y)=\sqrt{2}(y^2-y)$；

(3) $x^2-(2-2\sqrt{2})x+3-2\sqrt{2}=0$；

(4) $(t+6)(t-6)=2(t-3)$；

(5) $(x+2)^2+(x-1)^2=x^2+6$；

(6) $2(x+5)(x-5)=(x-6)^2$。

16. 已知二次方程 $x^2-mx+2m=0$ 的一個根是 1。

(1) 求 m 的值；　　(2) 求另一個根。

17. 不須解方程，求方程 $2x^2-7x+2=0$ 的兩根之倒數的和。

18. (1) 解分式方程與根式方程時，為什麼驗根是必要的步驟？

(2) 解下列方程：

(i) $\dfrac{x-1}{x^2-2x}+\dfrac{x-2}{x^2-x}-\dfrac{x}{x^2-3x+2}=0$；

(ii) $\left(\dfrac{x^2-1}{x}\right)^2+\dfrac{7}{6}\left(\dfrac{x^2-1}{x}\right)-4=0$；

(iii) $x^2+3x-\dfrac{20}{x^2+3x}=8$；

(iv) $\sqrt{5x+4}-\sqrt{2x-1}-\sqrt{3x+1}=0$；

(v) $x^2+3x-2\sqrt{x^2+3x-1}-4=0$；

(vi) $x^4-25x^2+144=0$；

(vii) $(x^2+5x)^2-2x^2-10x-24=0$；

(viii) $(x^2+5x-12)(x^2+5x+2)=32$。

19. 解下列方程組：

(1) $\begin{cases}(x+y)^2-3(x+y)=54\\(x-y)^2-5(x-y)=14\end{cases}$；　(2) $\begin{cases}a+aq^2=15\\aq+aq^3=30\end{cases}$；

(3) $\begin{cases}\dfrac{3x-5y}{2xy}=\dfrac{4}{3}\\\dfrac{3}{x}+\dfrac{1}{y}=4\end{cases}$；　(4) $\begin{cases}\sqrt{\dfrac{x}{y}}+\sqrt{\dfrac{x}{y}}=\dfrac{5}{2}\\x+y=5\end{cases}$；

(5) $\begin{cases}x+xy+y=11\\x^2y+xy^2=30\end{cases}$。

20. (1) 從$\begin{cases}mu=s(m-1)\\f=s-u\end{cases}$ (這裡 $m\neq 0$)導出 f、s、m 間的一個關係式；

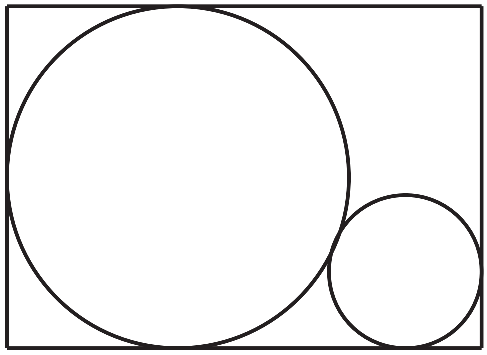
(2) 從$\begin{cases}I_k=I_a+I_c\\I_c=mI_k\end{cases}$ (這裡 $m\neq 1$)導出 I_c 、m、 I_a 間的一個關係式。

21. 甲蓄水池內有水 380 m^3，乙蓄水池內有水 1500 m^3。甲蓄水池每小時進水 80 m^3，乙蓄水池每小時抽出水 60 m^3，。經過幾小時，兩個蓄水池的水一樣多？

22. 甲、乙二人在 400 m 長的環形跑道上練習自行車，甲的速度比乙的速度快。當他們都從某處同時出發背向行駛時，每隔 40 秒鐘就相遇一次，同向行駛時，每隔 6 分 40 秒鐘相遇一次。求甲、乙兩人騎車的速度。

23. 某項工作，甲、乙兩人合作，16 天可以完成。如果共同作了 4 天後，剩下的工作由乙單獨完成，需要的時間比由甲一人完成全部工作的時間多 12 天。問單獨完成全部工作，甲、乙各需要多少天？

24. 如圖，一塊長 25 cm，寬 18 cm 的鐵板，在剪掉一個與三邊相切的圓後，剩下的鐵板能剪出之最大圓的直徑是多少？



(第 24 題)

25. 某工廠每天能生產甲種零件 500 個，或者乙種零件 600 個，或者丙種零件 750 個。甲、乙、丙三種零件各一個才能配成一套。現在要在 30 天內生產最多的成套產品，甲、乙、丙三種零件各應生產多少天？

26. 一種殺蟲劑 40 kg，含藥率是 15%，現在要用含藥率較高的同樣殺蟲劑 50 kg 與它混合，使混合後的含藥率在 25%與 30%之間(不包括 25%與 30%)，求所用殺蟲劑的含藥率。

27. 解下列不等式或不等式組，並在數軸上表示出它們的解集：

(1) $2(x+1)+\frac{x-2}{3}>\frac{7x}{2}-1$；

(2) $2x-\frac{x-1}{2}<\frac{2x-1}{3}+\frac{x+1}{6}$；

(3) $\frac{1}{2}\{1-2[x-4(x-1)]\}\le 4x$；

(4) $-1\le\frac{3-x}{2}\le 2$；

(5) $-4\le\frac{3x-5}{2}\le -2$；

(6) $\begin{cases}3x-5\ge 2x+3\\ \frac{x-5}{2}\le 3x+1\end{cases}$。

28. 在數軸上表示適合下列條件的數之集合：

(1) 使不等式 $x\le 5$ 或 $x>-2$ 成立；

(2) 使不等式 $x\le 5$ 與 $x>-2$ 同時成立；

(3) 使不等式 $x>5$ 或 $x>-2$ 成立；
(4) 使不等式 $x>5$ 與 $x>-2$ 同時成立；
(5) 使不等式 $x>5$ 或 $x\le -2$ 成立；
(6) 使不等式 $x>5$ 與 $x\le -2$ 同時成立。

29. 計算：

(1) $(-3x^{\frac{1}{3}}y^{\frac{1}{4}})(2x^{-\frac{2}{3}}y^{-\frac{1}{2}})(-x^{-\frac{2}{3}}y^{\frac{1}{4}})$；

(2) $\left(\dfrac{m^4n^{-4}}{m^{-1}n}\right)^{-3}\div\left(\dfrac{m^{-2}n^2}{mn^{-1}}\right)^5$；

(3) $(3x^{\frac{1}{2}}-2y^{-\frac{1}{2}})(3x^{\frac{1}{2}}+2y^{-\frac{1}{2}})$；

(4) $\sqrt[3]{3}\times\sqrt[4]{4}\times\sqrt[6]{6}$；

(5) $\sqrt[3]{x^{-5}y^2\sqrt{x^3y}}$；

(6) $(\sqrt{2x}-3\sqrt[4]{y})(\sqrt{2x}+3\sqrt[4]{y})$。

30. (1) 為什麼利用對數可以用加法計算乘法，用減法計算除法？
(2) 用科學記數法寫出下列各數，並說明它們的常用對數之首數：

54890、0.08351、4.903、213.7。

31. 已知 $\log 2=0.3010$，試判定 $2^7\times 8^{11}\times 5^{11}$ 是幾位數。。

32. 利用對數計算：

(1) $\sqrt[8]{\dfrac{0.3586\times\sqrt[4]{257}}{0.00501}}$；　　(2) $\dfrac{32.85^2-12.62^2}{32.85\times 12.64}\div\sqrt{\dfrac{32.85}{12.64}}$。

33. 求下列函數的自變量之取值範圍：

(1) $y=\dfrac{x-5}{x^2-3x+2}$；　　(2) $y=\sqrt{8-2x-x^2}$。

34. 已知 $y_1 = 2x - 3$ 、 $y_2 = -3x + 7$ ，計算 x 取何值時，

(1) $y_1 > y_2$ ；　(2) $y_1 = y_2$ ；　(3) $y_1 < y_2$ ；

並用圖形加以說明。

35. 計算 k 為何值時，函數 $y = -x^2 + 2x + k$ 的圖形與 x 軸

(1) 相交於一點；　(2) 相交於兩點；　(3) 不相交；

並用圖形加以說明。

36. 函數 $y = x^2 + px + q$ 的最小值是 4。在 $x = 2$ 時， $y = 5$ 。求 p、q 的值。

37. 設 α 是 $0^\circ \sim 180^\circ$ 間的角。根據下列條件求 α 。如果有解，求出所有的解：

(1) $\tan\alpha = -1$　(2) $\cos\alpha = 0$　(3) $\sin\alpha = \frac{1}{2}$

(4) $\cot\alpha = \sqrt{3}$　(5) $\sin\alpha = -\frac{1}{2}$　(6) $\cos\alpha = \sqrt{2}$

38. 直角三角形 ABC 之斜邊 AB 上的高 $CD = 21$ cm，$AD = 18$ cm，解這個三角形(邊長保留兩位有效數字，角度精確到1°)。

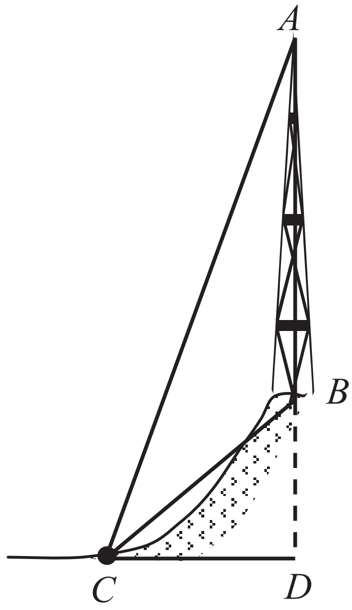
39. 在半徑為 10.0 cm 的圓中，作內接正五角星 $ABCDE$，求這五角星的頂點 A 與 B 間之距離，以及頂點 A 與 C 間的距離(精確到 0.1 cm)。

40. 已知 $B = 30^\circ$ 、 $c = 150$ 、 $b = 50\sqrt{3}$ ，滿足條件的三角形 ABC 有什麼特點？

41. 求證：在△ABC 中，如果 $\sin^2 A + \sin^2 B = \sin^2 C$ ，那麼這個三角形一定是直角三角形。

42. 平行四邊形 $ABCD$ 中，$AB = 8$、$AD = 5$、 $\angle A = 60^\circ$ ，如果取 A 作原點，AB 所在的直線作 x 軸，C 點所在的象限作第一象限，求它的各個頂點之座標。

43. 如圖，在小山頂上有一電視發射塔，塔高 AB 是 50 m，在平地 C 點，測得 B 的仰角是 $40°$、A 的仰角是 $70°$，求小山 BD 的高(精確到 1 m)。



(第 43 題)

44. 設 l 是第一、第三象限內兩座標軸夾角的平分線，在 l 上求一點，使它與兩點 $A(8, 0)$、$B(1,-3)$的距離相等。

45. 求證：$(-4, 3)$、$(8, 8)$、$(13,-4)$、$(1,-9)$是一個正方形的四個頂點。

46. 已知在 n 個數據中，x_1 出現 f_1 次，x_2 出現 f_2 次，…，x_k 出現 f_k 次($f_1+f_2+\cdots+f_k=n$)，$\overline{x}$ 是這 n 個數據的平均數。求證：

$$f_1(x_1-\overline{x})+f_2(x_2-\overline{x})+\cdots+f_k(x_k-\overline{x})=0$$ 。

47. 甲、乙兩種水稻，測得每種水稻各 10 株高度如下(單位：cm)：

甲：	12	9	16	18	14	8	12	10	17	11
乙：	12	15	14	16	15	13	12	10	12	10

哪種水稻的高度比較整齊？

48. 在生產過程中，測得特麗龍的纖度(表示纖維粗細的一種量)有如下的 100 個數據，試列出樣本的頻率分布表與繪出頻率分布直方圖。

1.36	1.49	1.43	1.41	1.37	1.40	1.32	1.42	1.47	1.39
1.41	1.36	1.40	1.34	1.42	1.42	1.45	1.35	1.42	1.39
1.44	1.42	1.39	1.42	1.42	1.30	1.34	1.42	1.37	1.36
1.37	1.34	1.37	1.37	1.44	1.45	1.32	1.48	1.40	1.45
1.39	1.46	1.39	1.53	1.36	1.48	1.40	1.39	1.38	1.40
1.36	1.45	1.50	1.43	1.38	1.43	1.41	1.48	1.39	1.45
1.37	1.37	1.39	1.45	1.31	1.41	1.44	1.44	1.42	1.47
1.35	1.36	1.39	1.40	1.38	1.35	1.42	1.43	1.42	1.42
1.42	1.40	1.41	1.37	1.46	136	1.37	1.27	1.37	1.38
1.42	1.34	1.43	1.42	1.41	1.41	1.44	1.48	1.55	1.37