數學 2

進階講義

標準差

景美女中．莊嘉銘 老師

信望愛文教基金會

7-1-4　標準差

定理證明或說明

1. 前言

一般而言，統計上經常以一簡單的數值，如平均數等，來代表整個母群體的中心趨勢，並藉以做為統計分析比較的衡量標準。可是由於母群體中之個體彼此互有差異，只用一個數值來代表整個母群體的狀況，有時過分的含糊與籠統，為彌補平均數的不足，還須另覓一簡單的數值，即差量，以表示母群體內個體彼此之間分散的情形，並藉以測量平均數的可靠程度。其特性為差量大，則母群體內各數值分散必大；反之差量小，則母群體內各數值必然較密集。依據這一特性，我們可以瞭解，倘若某一母群體的差量較小，則以平均數代表整個母群體的中心趨勢是極確實而可靠的，反之則否。

2. 變異數

設有 n 個資料 $x_1, x_2, \cdots, x_n$，則此資料的變異數是所有資料的離差平方之平均，

即 $\sigma^2 = \dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$，其中 μ 是母體平均數

3. 標準差

設有 n 個資料 $x_1, x_2, \cdots, x_n$，則此資料的標準差是變異數的開方，

即 $\sigma = \sqrt{\dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}}$

4. 平移

若 $y_i = ax_i + b$，其中 $i = 1, 2, \cdots, n$

則 $\sigma_y = |a|\sigma_x$

關鍵字

變異數，標準差

例題 1

求下列5個數據的變異數和標準差

$111, 112, 113, 114, 115$

Ans：

先求出平均數 $\mu_x = \dfrac{111+112+113+114+115}{5} = 113$，

則變異數 σ^2

$$= \frac{(111-113)^2+(112-113)^2+(113-113)^2+(114-113)^2+(115-113)^2}{5}$$

$= 2$

及標準差 $\sigma = \sqrt{2}$

例題 2

下表為某公司40名員工的薪資(萬元)次數分配表，求薪資的標準差

數值	0～2	2～4	4～6	6～8
次數	7	14	11	8

Ans：

這40個數的平均為 $\mu = \dfrac{1\times7+3\times14+5\times11+7\times8}{40} = \dfrac{160}{40} = 4$

組中點 x_i	次數	(離均差)^2
1	7	$(1-4)^2 = 9$
3	14	$(3-4)^2 = 1$
5	11	$(5-4)^2 = 1$
7	8	$(7-4)^2 = 9$

得標準差為 $\sigma = \sqrt{\dfrac{1}{40}(9\times7+1\times14+1\times11+9\times8)} = \sqrt{\dfrac{160}{40}} = 2$

例題 3

某次考試，某班的數學成績不太理想，全班 30 位成績的算術平均數為 36 分，標準差為 12 分，全班最高也僅 66 分。該班數學老師決定將每位學生的原始成績 x 採取線型函數 $y=ax+b$ 調整，並設定 y 成績的最高分為 100 分，y 成績的算術平均數為 60 分，作為成績的正式紀錄，則 y 成績的標準差為何？

Ans：

由題意可知 $\mu_x=36,\sigma_x=12$

因為 $y=ax+b \Rightarrow \begin{cases} 60=a\times 36+b \cdots\cdots ① \\ 100=a\times 66+b \cdots\cdots ② \end{cases}$

由①② $\Rightarrow 30a=40 \Rightarrow a=\dfrac{4}{3} \Rightarrow b=12$

即調整的公式為 $y=\dfrac{4}{3}x+12$

$\therefore \sigma_y=\dfrac{4}{3}\sigma_x=\dfrac{4}{3}\times 12=16$

例題 4

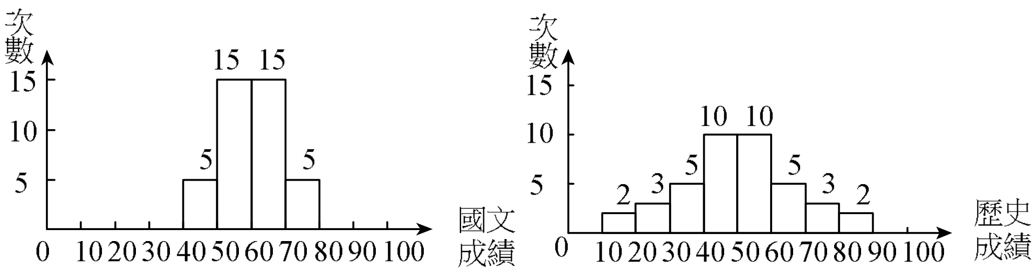
某班級國文、歷史成績分布情況直方圖依序如下面左右兩圖，根據該圖下列哪些推論是合理的？

(1) 國文分數的全距比歷史分數的全距大

(2) 國文分數的算術平均數比歷史分數的算術平均數低

(3) 國文分數的中位數比歷史分數的中位數大

(4) 國文分數的標準差比歷史分數的標準差小



Ans：

(1) 國文全距：$75-45=30$，歷史全距：$85-15=70$，所以歷史全距較大

(2) 國文平均：60，歷史平均：50，國文平均較高

(3) 國文中位數：60，歷史中位數：50，國文中位數較大

(4) 如圖所示，國文分散程度較歷史小，即歷史標準差>國文標準差

故選(3)(4)

例題 5

高二甲班 50 名學生的期中考數學成績，中位數 74 分，算術平均數 75.2 分，眾數 76 分。今發現某生成績應為 76 分誤登記為 86 分，試問下列哪些統計量不變？

(1) 算術平均數 (2) 中位數 (3) 全距 (4) 標準差 (5) 眾數

Ans：

(1) 算術平均數應修正為 $75.2-\dfrac{86-76}{50}=75$ 分

(2) 86 分與 76 分都大於中位數 74，位於中位數的同一側，且因眾數為 76 分，76 分不會成為從小到大唯一的第 26 個數，故中位數不變

(3) 若 86 分為全班最高分，全距有可能改變

(4) 原本的標準差為 $\sqrt{\dfrac{1}{50}(\sum_{i=1}^{49}x_i^2+86^2)-(75.2)^2}=\sqrt{\dfrac{1}{50}(\sum_{i=1}^{49}x_i^2)-5507.12}$ 而後來的標準差為 $\sqrt{\dfrac{1}{50}(\sum_{i=1}^{49}x_i^2+76^2)-(75)^2}=\sqrt{\dfrac{1}{50}(\sum_{i=1}^{49}x_i^2)-5509.48}$ 相較之下，標準差變小 .

(5) 原眾數為 76 分，修正後仍為 76 分

故答案為(2)(5)

溫故知新

習題 1

求數據 101 到 111 的標準差

習題 2

某班50人的數學成績，整理如下表：

分數	30～40	40～50	50～60	60～70	70～80
人數	5	10	10	20	5

求數學成績的標準差為何？

習題 3

設數據X：$x_1, x_2, \cdots, x_n$，數據Y：$y_1, y_2, \cdots, y_n$，且$y_i = 3x_i + 15$，已知數據X的標準差為5，求數據Y的標準差為何？

習題 4

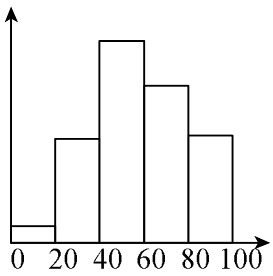
下列甲、乙、丙、丁、戊五個圖形，是某校高三學生參加學測各科成績的次數分配表的直方圖，其縱坐標表人數，橫坐標表成績，則下列敘述何者必正確?

(1)算術平均數：國文＞數學

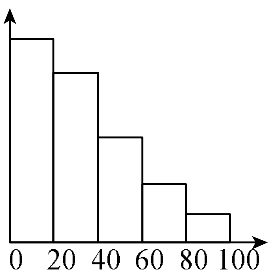
(2)標準差：英文＝數學

(3)中位數：數學＞社會

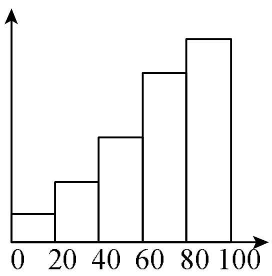
(4)眾數：英文＞自然



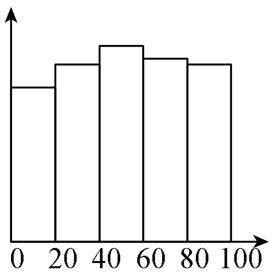
國文



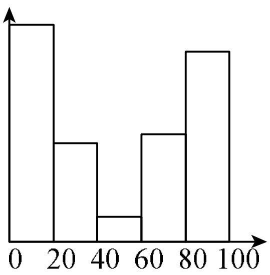
英文



數學



自然



社會

習題 5

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？

(1) $\sigma_B = 2\sigma_A$　(2) $\sigma_C = \sigma_A$　(3) $\sigma_C > \sigma_A$　(4) $\sigma_D > \sigma_A$　(5) $\sigma_D = (\sigma_A)^2$

習題 6　【87 年自然組】

某年聯考甲、乙兩科成績的直方圖如下圖所示(由於考生人數眾多，成績分布的直方圖可視為平滑的曲線)，則下列那些敘述是正確的？

(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

習題 7　【88 年推甄】

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$3,6,1,5,4,8,6,7,5$，問下列選項，何者為真？

(1) 新數據的算術平均數為5　(2) 新數據的標準差為2

(3) 原數據的算術平均數為2.45　(4) 原數據的標準差為0.2

(5) 原數據的中位數為2.45

習題 8 【指考乙 97】

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，請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。)

(1) 0.240毫米
(2) 1.61毫米
(3) 6.10毫米
(4) 155毫米
(5) 1041毫米

解答與解析

習題 1：$\sqrt{10}$

【詳解】先求出平均數$\mu_x=\dfrac{101+102+\cdots+111}{11}=106$，

則變異數σ^2

$$=\frac{(101-106)^2+(102-106)^2+\cdots+(110-106)^2+(111-106)^2}{11}$$

$=10$

$\therefore$標準差$\sigma=\sqrt{2}$

習題 2：$\sqrt{136}$

【詳解】先求出平均數

$$\mu_x=\frac{35\times5+45\times10+55\times10+65\times20+75\times5}{50}=\frac{2850}{50}=57$$

組中點 x_i	次數	(離均差)^2
35	5	$(35-57)^2=484$
45	10	$(45-57)^2=144$
55	10	$(55-57)^2=4$
65	20	$(65-57)^2=64$
75	5	$(75-57)^2=324$

得標準差為

$$\sigma=\sqrt{\frac{1}{50}(484\times5+144\times10+4\times10+64\times20+324\times5)}$$

$$=\sqrt{\frac{6800}{50}}=\sqrt{136}$$

習題 3：15

【詳解】若 $y_i=ax_i+b$，則 $\sigma_y=|a|\times\sigma_x$

故所求為 $\sigma_y=3\sigma_x=3\times5=15$

習題 4：(2)(3)

【詳解】(1) 國文平均較接近於 60，數學平均較接近 80，故國文 $<$ 數學

(2) 正確

(3) 數學的中位數介於 $60\sim80$ 分，社會的中位數介於 $20\sim40$ 分

(4) 英文眾數介於 $0\sim20$ 分，自然眾數介於 $40\sim60$ 分

故選(2)(3)

習題 5：(1)(2)(4)

【詳解】從題意可知 $B=2A$，$C=A+995$，$D=A^2$，

且若 $y_i=ax_i+b$，則 $\sigma_y=|a|\times\sigma_x$

(1) ○： $\sigma_B=2\sigma_A$

(2) ○： $\sigma_C=\sigma_A$

(3) ╳

(4) ○： $\because D$ 的數據較分散 $\therefore\sigma_D>\sigma_A$

(5) ╳

故選(1)(2)(4)

習題 6：(C)(D)

【詳解】由於甲、乙的成績分布的圖形近似對稱故算術平均數、中位數及眾數大約在中央的位置，因此甲的算術平均數、中位數及眾數均小於乙又甲的圖形比乙來得分散，故甲的全距與標準差均大於乙

故選(C)(D)

習題 7：(1)(2)(3)(5)

【詳解】新數據的平均數為$\dfrac{3+6+1+5+4+8+6+7+5}{9}=\dfrac{45}{9}=5$

新數據的標準差為$\sqrt{\dfrac{(3-5)^2+(1-5)^2+\cdots+(7-5)^2+(5-5)^2}{9}}=\sqrt{4}=2$

又新數據為原數據乘以100，再減去240

所以原數據的平均數為$\dfrac{(5+240)}{100}=2.45$

原數據的標準差為$\dfrac{2}{100}=0.02$

且中位數為由小到大排後第5個數據$\Rightarrow 2.45$

故選(1)(2)(3)(5)

習題 8：(4)

【詳解】標準差$=6.1$吋$=6.1\times 25.4$毫米≈ 155毫米