

關於標準分和正態分佈的教學討論

梁子傑
退休中學教師

標準分

若一組數據 $\{x_1, x_2, \dots, x_i, \dots, x_n\}$ 的平均值 (mean) 為 μ ，標準差 (standard deviation) 為 σ ，則我們定義數據 x_i 的標準分 (standard score) z_i 為 $\frac{x_i - \mu}{\sigma}$ 。在一般的數學教科書中，在這個定義之後，通常都會引入以下類似的例子或練習：

1. 在某次測驗中，一班學生得到的平均分和標準差分別為 54 分和 6 分。如果小吉在這次測驗中的得分是 60 分，求他的標準分（答案為 1）。
2. 志健和全班同學的平均體重為 68 kg。若志健的體重為 77 kg，標準分為 0.6，求全班同學體重的標準差（答案為 15 kg）。
3. 在一個智力測驗中，得分的標準差為 15 分，而梓奇的個人得分為 135 分，他的標準分為 1.6。求該智力測驗的平均分（答案為 111 分）。
4. 在一項比賽中，芷潔所得的標準分為 -1.8。若所有參賽者的平均得分和標準差分別為 269 分和 15 分，求芷潔的實際得分（答案為 242 分）。

以上 4 題可分為兩類：第 1 題是第一類，只要把已知數值直接代入公式，便可求得答案；第 2 至第 4 題是另一類，在解題時須用上解方程的技巧，做一點「移項」，才可得到所求的答案。無論如何，這些題目的難度應該不算太大。

在近年數學科文憑試的試卷中，每年都總會有一至兩條題目與標準分有關，而它們的難度卻比一般教科書的習題為高。先討論以下三條模仿¹近年考試的題目。

¹ 使用模仿題目，除了避免版權的問題之外，還有一個重要原因，下文會有解說。

5. 在一次考試中，考試得分的平均值為 45 分。某男生在該次考試得到 27 分且他的標準分為 -1.2 。若某女生在該次考試的標準分為 1.6，求她的考試得分。

6. 下表顯示三名學生在某次考試的得分及其對應的標準分。

得分	58 分	x 分	86 分
標準分	-1.5	1	2

求 x 。

7. 在一次測驗，某兩名學生測驗得分之差及標準分之差分別為 21 分及 1.4。求該測驗得分的標準差。

由定義可知，若要求得第 5 題中該女生的考試得分，我們必須知道該考試的平均分、標準差和該女生的標準分。我們可以從已知條件直接得到平均分和標準分，但標準差卻沒有提供。幸好，我們知道男生的得分和標準分，因此可以透過設標準差為 σ 分和建立方程 $\frac{27-45}{\sigma} = -1.2$ ，解得 $\sigma = 15$ 。最後設該女生的得分為 x 分和建立方程 $\frac{x-45}{15} = 1.6$ ，便可求得該女生的考試得分為 69 分了。

第 6 題則較複雜。這次題目連平均值也沒有提供。那麼，我們須先分別設考試的平均分和標準分為 μ 分和 σ 分，與及建立兩條方程，即

$$\frac{58-\mu}{\sigma} = -1.5 \text{ 和 } \frac{86-\mu}{\sigma} = 2,$$

然後解該聯立方程組，得到 $\mu = 70$ 和 $\sigma = 8$ ，最後再解方程 $\frac{x-70}{8} = 1$ ，便可知 $x = \underline{78}$ 。

第 7 題的難度又再提升。給定的條件連實際的得分和標準差與及平均值都沒有。這時候我們要假設兩名學生的測驗分分別為 x_1 和 x_2 ，平均分和標準分分別為 μ 和 σ ，那麼我們可以建立以下方程：

$$\frac{x_1 - \mu}{\sigma} - \frac{x_2 - \mu}{\sigma} = 1.4$$

化簡得

$$x_1 - \mu - x_2 + \mu = 1.4 \sigma$$

$$1.4 \sigma = x_1 - x_2 = 21$$

即

$$\sigma = \frac{21}{1.4} = 15$$

∴ 該測驗的標準差為 15 分。

可是，細想想：第 7 題是否真的很複雜呢？上面解題的最後一步，是把 21 分除以標準分 1.4，這豈不是等同於計算 1 個標準分對應於多少分數嗎？（情況就好像知道 3 個橙值 18 元，求 1 個橙值多少錢，又或者知道 1.5 塊薄餅值 72 元，求 1 塊薄餅值多少錢一樣。）而這個商，又不是正好等於標準差嗎？

這個做法又令我想到了，其實第 6 題可以有一個快捷的方法來完成。首先，我知道標準差 $\sigma = \frac{86 - 58}{2 - (-1.5)} = 8$ 。然後，因為 x 比 86 分低 1 個標準差，所以 $x = 86 - 8 = \underline{78}$ 。

雖然上述方法能夠獲得正確的答案，但是卻稍欠了一點說服力。回看標準分的定義，我們有 $z = \frac{x - \mu}{\sigma}$ 。把分母分開，這定義可變換成

$$z = \frac{1}{\sigma}x - \frac{\mu}{\sigma}。$$

留意： μ 和 σ 都是常數；換言之， $\frac{1}{\sigma}$ 和 $-\frac{\mu}{\sigma}$ 都是常數。由此得知，上面的寫法與一條直線方程 $z = mx + c$ 毫無分別，即 z 與 x 之間是一個線性關係！再換句話說，我們可以把第 6 題中的表視之為列出三個共線點的坐標！因此，我們可以利用比較直線斜率的方法來解第 6 題。方法如下：

$$\frac{x - 58}{1 - (-1.5)} = \frac{86 - 58}{2 - (-1.5)} = 8$$

$$x - 58 = 2.5 \times 8 = 20$$

∴

$$x = \underline{78}$$

至於第 5 題，男生和女生的「坐標」是清楚的（我們可以設該女生的得分為 x ）。留意：平均分的標準分是 0。換言之，我們同樣有三個共線點的坐標，因此，也可以利用求斜率的方法來求女生的得分，即：

$$\frac{x-45}{1.6-0} = \frac{27-45}{-1.2-0} = 15$$

$$x-45 = 1.6 \times 15 = 24$$

$$\therefore x = \underline{69}$$

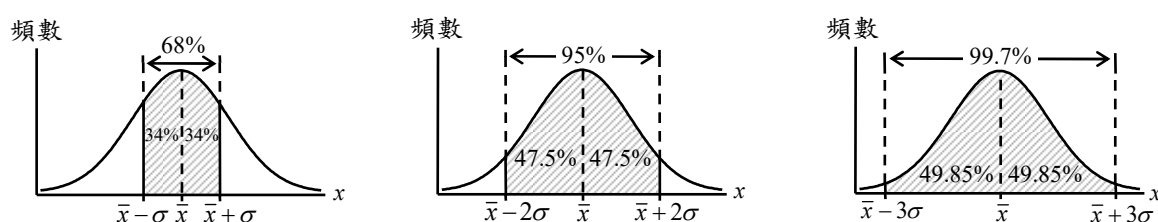
至於標準分的教學之中，還有一個部分是需要注意的，可惜一般教科書都甚少提及，那就是數據改變對標準分的影響。例如：

8. 在一次測驗之後，老師將全班每名同學的得分都提升了 50%，然後再把每人的得分增加 8 分。若致杰原本所得的標準分為 -2，問經過分數調整之後，他的標準分將會是多少？

要知道，把每個數據乘以某個常數，又或者把每個數據加上某個常數，對平均值都有同樣的影響，但只有乘除會影響標準差。然而，在標準分的定義中，分子是 x 與 μ 的相差，換言之，加數的影響會被定義中的減號所消去，而分子、分母之間共同倍數又會在除法中化簡，最後的標準分是保持不變的。因此，經過分數調整後的標準分仍為 -2。

正態分佈

在日常生活之中，很多數據都會成正態分佈。而在數學科必修部分中，學生須了解數據分佈的百分比與正態分佈有下圖中的關係：



由此，學生須懂得完成以下的問題：

9. 在一個測試中，學生的智商成正態分佈，其中他們智商的平均值為 124，標準差為 9。求智商介乎於 115 和 142 之間的學生人數百分比。

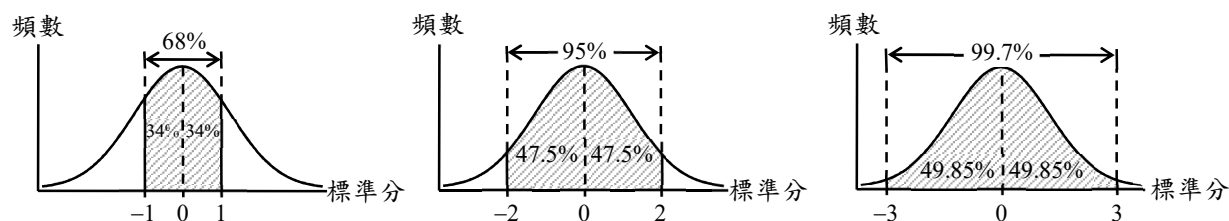
通常教科書的解答會以以下的方式呈現：

$$\begin{aligned}
 &\text{由於} \quad 115 = 124 - 9 = \mu - \sigma \\
 &\text{且} \quad 142 = 124 + 2 \times 9 = \mu + 2\sigma, \\
 &\text{因此} \quad \text{百分比} = 34\% + 47.5\% \\
 &\quad \quad \quad = \underline{81.5\%}
 \end{aligned}$$

對於很多學生來說，上述的計算已經足夠清楚易明，毋須老師多作解釋。不過，過往經驗總告訴我，每年總有一些學生不明白，為何我們知道可以把 115 分拆成 $124 - 9$ ，又把 142 分拆成 $124 + 2 \times 9$ 呢？

在數學延伸單元一（微積分與統計）中，當談及正態分佈時，有一個標準的方法來進行上述演算。其實我們可以借用這個方法向學生解釋有關運作。那就是計算數據的標準分。譬如：當 $x = 115$ 時，它對應的標準分為 $\frac{115-124}{9} = -1$ ；又當 $x = 142$ 時，它對應的標準分為 $\frac{142-124}{9} = 2$ 。這兩個數都與上述以分拆方法找到的結果相同。所以當談及正態分佈的應用題時，應該同樣使用標準分，好讓大多數學生都能掌握。

事實上，對於前面引入的正態分佈圖，我認為那些橫軸不應以原始數據 x 來表達，這會為能力較弱的學生帶來困擾。個人認為，為了讓生更容易明白和了解所使用的方法，我們應該將橫軸以標準分 z 來表示。



回看標準分

看過上面的正態分佈圖之後令我駭然想起一個事實：標準分不可能任意設定大小！

留意：數據 x 、平均值 μ 和標準差 σ 並非互相獨立的，它們之間的數值互相牽連，不可以任意設定，尤其是在正態分佈的情況下，標準分 $z = \frac{x - \mu}{\sigma}$ 能夠大於 3 或小於 3 的機會，是非常之小的（出現的概率只有

0.3%)。為了保持習題的合理性，因此我們在設定標準分的練習時，標準分的絕對值必須小於 3，甚至是 2。

前面提到，本文中所引用的第 5 至第 7 題，是模仿公開試試題而寫成的。我不直接使用那些題目，一來是避免涉及版權的爭議，二來就是原題中，那些標準分大得不合理。例如：2021 年題目中的兩個標準分分別等於 5 和 -7；2020 年提及有兩個標準分，它們的差竟然高達 6。我明白，那些題目是用來考核學生是否知道標準分的計算方法，但不合理的題目的確會令到一些懂得「慎思明辨」的學生猶豫，反而影響他們作答時的信心。除非我們同意，公開考試就是用來挑選那些只懂機械計算的學生，否則在設定這類題目時，應該注意有關數值的合理性。

作者電郵：jckleung@netvigator.com