3下 2-1 四分位數 Quartiles

當我們試圖以一個數字來概括一組資料時，通常會選擇使用平均值或中位數，但如果我們希望深入瞭解資料的分佈情況，這些統計量可能顯得不夠完整。

翻譯示例：
When we want to use a numerical value to represent a set of data, we usually use the mean or median, but if we want to further understand the distribution of the data, it may be insufficient.

3、3、4、5、5、6、7、7、7、8、8、9、11、11、11
平均數：7、中位數：7

2、3、4、4、7、7、7、7、8、8、8、8、10、11、11
平均數：7、中位數：7

這兩組資料的平均數與中位數都是 7，但數值的分布卻不同。為了進一步了解資料分布的情形，將各組資料分別由小到大排列，再分成四等分，此時便可用各組資料$\frac{1}{4}$、$\frac{2}{4}$、$\frac{3}{4}$位置的數值來了解資料分布的情形，上述位置的數值稱為**四分位數**。

翻譯示例：
The mean and median of both data sets are 7, but the distribution of the values is different. In order to further understand the distribution of the data, the data for each group is arranged from the least to the greatest and divided into four equal parts. Then, the values of the data at the $\frac{1}{4}$, $\frac{2}{4}$, and $\frac{3}{4}$ position of the data can be used to understand the distribution of the data. These values at these positions are called quartiles.

其中，資料的$\frac{1}{4}$位置，稱為**第 1 四分位數**，簡記為 Q_1；

資料的$\frac{2}{4}$位置，稱為**第 2 四分位數**，簡記為 Q_2，也就是中位數；

資料的$\frac{3}{4}$位置，稱為**第 3 四分位數**，簡記為 Q_3。

翻譯示例：

The $\frac{1}{4}$ position of the data is called the first quartile (lower quartile), abbreviated as Q1.

The $\frac{2}{4}$ position of the data is called the second quartile, abbreviated as Q2, which is the median.

The $\frac{3}{4}$ position of the data is called the third quartile (upper quartile), abbreviated as Q3.

無論哪一種取法，都是為了要使小於或等於中位數的資料至少占全部資料的$\frac{\mathbf{1}}{\mathbf{2}}$，且大於或等於中位數的資料也至少占全部資料的$\frac{\mathbf{1}}{\mathbf{2}}$。

翻譯示例：

Regardless of which method we use to find the median, we need to ensure that the data points with values less than or equal to the median account for at least $\frac{1}{2}$ of all the data points, and the data points with values greater than or equal to the median also account for at least $\frac{1}{2}$ of all the data points.

Review: How do you find the median of a data set? Order the data from least to greatest and find the middle number in the data set, or the mean of the two middle numbers.

在計算四分位數時，我們也用相同的方法。

第 1 四分位數（Q1）在整體資料中的位置，需符合兩個條件：

(1)小於或等於 Q1 的資料筆數至少占全部資料筆數的 $\mathbf{\frac{1}{4}}$，

(2)大於或等於 Q1 的資料筆數至少占全部資料筆數的 $\mathbf{\frac{3}{4}}$。

翻譯示例：

The first quartile (Q1) must meet two conditions in the overall data:

(1) The number of data points less than or equal to Q1 accounts for at least $\frac{1}{4}$ of all data points.

(2) The number of data points greater than or equal to Q1 accounts for at least $\frac{3}{4}$ of all data points.

第 3 四分位數（Q_3）在整體資料中的位置，需符合兩個條件：

(1)小於或等於 Q3 的資料筆數至少占全部資料筆數的 $\mathbf{\frac{3}{4}}$，

(2)大於或等於 Q3 的資料筆數至少占全部資料筆數的 $\mathbf{\frac{1}{4}}$。

翻譯示例：

For Q3 to be valid, it must meet two conditions in relation to the overall data:

(1) The number of data points that are less than or equal to Q3 must be at least $\frac{3}{4}$ of the total number of data points.

(2) The number of data points that are greater than or equal to Q3 must be at least $\frac{1}{4}$ of the total number of data points.

如果資料有 12 筆資料，那麼四分位數要怎麼求呢？

第 1 四分位數（Q_1）：$12\times\frac{1}{4}=3$（是整數），故取第 3、4 筆資料的平均為 Q_1。

因為在第 3 筆和第 4 筆之間的所有數字（包含第 3 筆和第 4 筆）都可以滿足第 1 四分位數（Q_1）在整體資料中的位置，需符合小於或等於 Q1 的資料筆數至少占全部資料筆數的$\mathbf{\frac{1}{4}}$；且大於或等於 Q1 的資料筆數至少占全部資料筆數的$\mathbf{\frac{3}{4}}$，所以我們取其平均來表示。

翻譯示例：

If there are 12 data points, how do we calculate quartiles?

Q1: $12\times\frac{1}{4}=3$ (an integer), so we take the average of the 3rd and 4th data points as Q1.

Since all numbers between the 3rd and 4th data points (including the 3rd and 4th) can satisfy the position of Q1 within the overall data set, which must have at least $\frac{1}{4}$ of the data points less than or equal to Q1 and at least $\frac{3}{4}$ of the data points greater than or equal to Q1, we take the average of the two to represent Q1.

至於第 2 四分位數與第 3 四分位數也是用同樣的方法。

第 2 四分位數（Q_2）：$12\times\frac{2}{4}=6$（是整數），故取第 6、7 筆資料的平均為 Q_2。

第 3 四分位數（Q_3）：$12\times\frac{3}{4}=9$（是整數），故取第 9、10 筆資料的平均為 Q_3。

翻譯示例：

The same method is used for finding the second and third quartiles.

The second quartile (Q2): $12\times\frac{2}{4}=6$ (an integer), so the average of the 6th and 7th data is taken as Q2.

The third quartile (Q3): $12\times\frac{3}{4}=9$ (an integer), so the average of the 9th and 10th data is taken as Q3.

【第 m 四分位數（Q_m）的算法】

將 n 筆資料由小到大排列，並計算 $n\times\frac{m}{4}$ 的值（$m=1$、2、3），

(1) 如果 $n\times\frac{m}{4}$ 不是整數，令 T 是大於 $n\times\frac{m}{4}$ 的最小正整數，
則 Q_m 是「第 T 筆資料」。

(2) 如果 $n\times\frac{m}{4}$ 是整數，令 $T=n\times\frac{m}{4}$，則 Q_m 是「第 T 筆資料與第（$T+1$）筆資料的平均數」。

補充：國外的方法

When finding the lower and upper quartiles, students would need to find the mean of the two middle values in each half of the data set.

The lower quartile is the mean of the two middle values of the lower half.
The upper quartile is the mean of the two middle values of the upper half.

Example 1:

Consider the prices for jackets: $120, $300, $540, $550, $690, $400, $850.

Goal: Find the median, the lower quartile, and the upper quartile of the prices.

A. Explain how to find the median, then find the median

   Order the prices from the least to the greatest: $120, $300, $400, $540, $550, $690, $850.

   Identify the middle number or find the mean of the two middle numbers.

   The median is $540.

B. The lower quartile is the median of the lower half of the data or data to the left of the median. What is the lower half of the data? What is the lower quartile?

   The lower half of the data is $120, $300, $400.

   The low quartile is $300.

C. The upper quartile is the median of the upper half of the data or data to the right of the median. What is the upper half of the data? What is the upper quartile?

   The upper half of the data is $550, $690, $850.

   The upper quartile is $690.

Example 2:

Consider the prices for jackets: $120, $300, $540, $550, $690, $400, $850, $540.

Goal: Find the median, the lower quartile, and the upper quartile of the prices.

A. Explain how to find the median, then find the median

   Order the prices from the least to the greatest: $120, $300, $400, $540, $540 $550, $690, $850.

   Identify the middle number or find the mean of the two middle numbers.

   The median is $540.

B. The lower quartile is the median of the lower half of the data or data to the left of the median. What is the lower half of the data? What is the lower quartile?

   The lower half of the data is $120, $300, $400, $540. (You can't ignore this 540.)

   The low quartile is $350.

C. The upper quartile is the median of the upper half of the data or data to the right of the median. What is the upper half of the data? What is the upper quartile?

   The upper half of the data is $540, $550, $690, $850. (You can't ignore this 540.)

   The upper quartile is $620.

參考資料來源

1. 110 國中數學翰林版課本
2. Into math advanced 1

☆老師們可以自己從中選擇以做出適合自己學生程度的學習單或是在課堂中適時補充這些英文。

製作者:康橋國際學校 陳怡伶