

## 16.2. 理解分佈域和四分位數間距的概念 (Understanding the Concepts of Range and Inter-quartile Range)

- 分佈域 (range) 係 “最大的數據 – 最細的數據 嘅結果” 。
    - 留意當中的 “-” 是 “減”，不是 “至” 。
      - ◆ 好多時我哋都會以為分佈域係 “幾多至幾多” 。
    - 除此之外，分佈域亦是有單位的。
  - 例子： 五個人嘅身體分別為 50kg, 60 kg, 65 kg, 67 kg, 80 kg 。
    - 以上數據嘅分佈域 =  $80 - 50 = 30$  kg
  - 留意分佈域只係講俾我哋知 “最大同最細數據嘅差距”，但並唔一定可以準確咁指數據嘅分散程度。
    - 例如 A 組數據有一萬個數據，當中有 9999 個 10，一個 10000 。
      - ◆ 咁 A 組嘅分佈域就 = 9990 。
    - 而 B 組數據都係有一萬個數據，當中有 1000 個 10、1000 個 20...1000 個 100 。
      - ◆ 咁 B 組嘅分佈域就 =  $100 - 10 = 90$  。
    - 雖然 A 組嘅分佈域大好多，但到底邊一組數據分散 D 呢？ 咁就真係見仁見智。
      - ◆ 為咗可以再準確 D 知道數據嘅分佈，數學家就提出 “四分位數間距” 嘅概念。
  - 四分位數間距入面嘅 “四分” 就係指
    - 先將 D 數據先排好大細，然後喺 25%、50%同 75%度切落去。咁咪變咗做四份囉。
    - 而 “四分位數” 就係 25%、50%同 75%嗰個位嘅數據嘅值 (連單位)，
      - ◆ “四分位數” 嘅記號係  $Q_1$ ， $Q_2$ ， $Q_3$  。
      - ◆  $Q_3$  亦叫作上四分位數 (upper-quartile)
      - ◆  $Q_1$  就叫作下四分位數 (lower-quartile)
      - ◆ 而  $Q_2$  就係 50%嘅數，亦即係 “中位數”
  - 根據定義， 四分位數間距 =  $Q_3 - Q_1$ 
    - 留意同分佈域一樣，如果 D 數據係有單位嘅，咁四分位數間距都要單位。
- ☆ 留意上面講咗要將 D 數據排列好 (通常由細到大咁排)。
- 例如如果數據排好後係： 1, 2, 3, 6, 7, 7, 8, 11, 100
    - ◆ 咁  $Q_1$ ， $Q_2$ ， $Q_3$  就係上面 underline 咗嘅數
    - ◆ 即四分位數間距 =  $8 - 3 = 5$
  - 另外大家亦即識得由 “頻數累積曲線圖” 入面睇返嘅  $Q_1$ ， $Q_2$ ， $Q_3$  嘅數值出嚟。
    - ◆ 其實 “頻數累積曲線圖” 已經同我哋排咗 D 數據。所以可以好似右面嘅例子咁睇返 D 數值出嚟。

