

## 16.4. 理解分組數據和不分組數據的標準差之概念 (Understanding the Concept of Standard Deviation for Both Grouped and Ungrouped Data Sets)

### 16.4.1. 不分組數據的標準差

- 唔知你記唔記得咩係“分組數據”同“不分組數據”呢？
  - 不分組數據就係好簡單將所有數據一個個咁列晒出嚟。
    - ◆ 亦有時會用“頻數”嚟指出有幾多個數嘅“值”係一樣。
    - ◆ 總之最緊要係我哋知道所有數據嘅正確值。
  - 分組數據就係會將數據分組，例如“1-10”、“11-20”。然後利用“頻數”嚟指出每個組別入面有幾多個數據。
    - ◆ 而喺分組數據入面，我哋唔會知道所有數據嘅正確值（只係知個範圍）。

- 喺度先講咗不分組數據的標準差（通常用符號  $\sigma$  代表）嘅公式：

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}}$$

- 標準差通常用符號  $\sigma$  代表（寫法係以順時針方向寫個 O 字後繼續畫多條尾咁）
- $n$  係數據嘅數量
- $x_1$ 、 $x_2$  等係個別數據嘅值
- $\bar{x}$  係所有數據嘅平均數（如果題目冇俾，咁你就先要由 D 數據計佢出嚟）

◇ 留意課程入面提到要我哋識“方差”（variance）依個術語。

- ◆ “方差”通常用  $s$  代表。
- ◆ 而“方差”定義係  $s = \sigma^2$

### 16.4.2. 分組數據的標準差

- 分組數據嘅標準差其實差唔多，只不過我哋要考慮埋
  - 頻數（ $f_1$ 、 $f_2$  等）代差每個分組內有幾多個數據
  - 要用“組中點”代表每個分組內所有數據嘅值（符號都係  $x_1$ 、 $x_2$ ）。
  - 因此條公式就會係

$$\sigma = \sqrt{\frac{f_1(x_1 - \bar{x})^2 + f_2(x_2 - \bar{x})^2 + \cdots + f_n(x_n - \bar{x})^2}{n}}$$

- 知道咗條上面條咁煩、咁長嘅公式之後，我諗係時候講吓“標準差之概念”。
- 喺後面嗰頁，我哋嘗試吓去理解數學家係點諗到用上面條式。

## 16.4.3. 標準差之概念

- 首先，“標準差”可以睇成係“數據同標準嘅差距”（依個解釋係我作嘅）。
  - 所以先要定出用咩做標準。而標準差就用咗“數據嘅平均數”做標準。
  - 至於“差距”，最簡單就梗係“擺個別數據同平均數嘅差距”。
- 而因為唔同數據組可以有唔同數目嘅數據，所以為求公平，當我哋計“數據同平均數嘅差距”嘅時候，我哋要計“數據同平均數嘅差距”嘅平均數。
- 基於以上概念，我哋第一個建議用作標準差嘅公式係：

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

- 假如數據係 5,5,5（平均數 = 5）。
  - ◆ 而用上面條公式計一計，標準差 =  $[(5 - 5) + (5 - 5) + (5 - 5)] / 3 = [0 + 0 + 0] / 3 = 0$ 。
  - ◆ 咁都好正常，因為所有數據同平均數嘅差距都係零！
- 但假如數據係 2,5,8 呢？（2,5,8 嘅平均數都係 5）
  - ◆ 標準差 =  $[(2 - 5) + (5 - 5) + (8 - 5)] / 3 = [-3 + 0 + 3] / 3 = 0 / 3 = 0$
  - ◆ 咁又唔合理！所以第一個版本嘅公式設計上有問題。
- 好明顯我哋計到零嘅原因係因為 D “差距”有正有負，所以 cancel 咗。
  - 為咗令所有嘅差距變做正數，我哋不如將每個差距擺二次方。
  - 所以我哋將上面嘅建議改成：

$$\sigma = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}$$

- 雖然咁就解決咗“正、負差距互相 cancel”嘅問題，但因為擺咗二次方，所以好似將 D 差距“放大咗”。
  - 例如 8 同平均數 5 嘅差距只係 3，但用  $(x_1 - \bar{x})^2$  就計到 9。
- 為咗解決“心理上過意唔去”，所以就將個結果“開返方”。
- 因此，最終被接納嘅標準差公式係：

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}}$$

☆ 希望上面所作嘅故仔/解釋可以幫到大家明白同記條公式。