



大數據資料處理實務

李水彬

2023-09-01

Chapter 06-3 統計量-分散趨勢量數

分散趨勢量數

描述這組資料內容的 **差異程度**。

- 方法一: 每一個數與集中趨勢量數 (中位數或平均值) 的距離。
 - 變異數 Variance、標準差 Standard deviation、平均絕對離差 Mean absolute deviation
- 方法二: 位置參數間的距離代表全體資料的差異程度。
 - 全距、四分位距、四分位差。

方法一

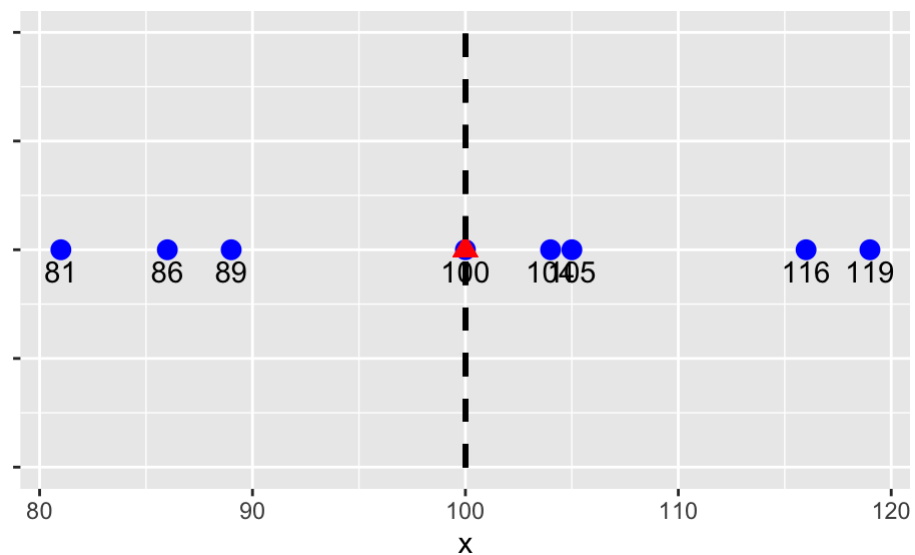
資料: 81, 86, 105, 119, 116, 100, 104, 89

排序: 81, 86, 89, 100, 104, 105, 116, 119

樣本平均值: $\frac{\sum_{i=1}^8 x_i}{8} = \frac{800}{8} = 100$

方法一

- 每個數與平均值的離差平方。
- 差越遠，離差平方越大。



| x | 離差 | 離差平方 |
|-----|-----|------|
| 81 | -19 | 361 |
| 86 | -14 | 196 |
| 89 | -11 | 121 |
| 100 | 0 | 0 |
| 104 | 4 | 16 |
| 105 | 5 | 25 |
| 116 | 16 | 256 |
| 119 | 19 | 361 |

變異數 & 標準差

數值

$$x_1, x_2, \dots, x_n$$

- 樣本變異數

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

-樣本標準差

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

變異數 & 標準差

- 數據

```
## [1] 81 86 89 100 104 105 116 119
```

```
mx<-mean(x$x)#平均值  
dd<-(x$x-mx)^2 # 離差平方  
mx;dd
```

```
## [1] 100
```

```
## [1] 361 196 121 0 16 25 256 361
```

離差平方和 = $361 + 196 + 121 + 0 + 16 + 25 + 256 + 361 = 1336$

變異數 = $\frac{1336}{8-1} = 190.857$

使用函數求變異數與標準差

```
var(x$x)#變異數
```

```
## [1] 190.8571
```

```
sd(x$x)#標準差
```

```
## [1] 13.81511
```

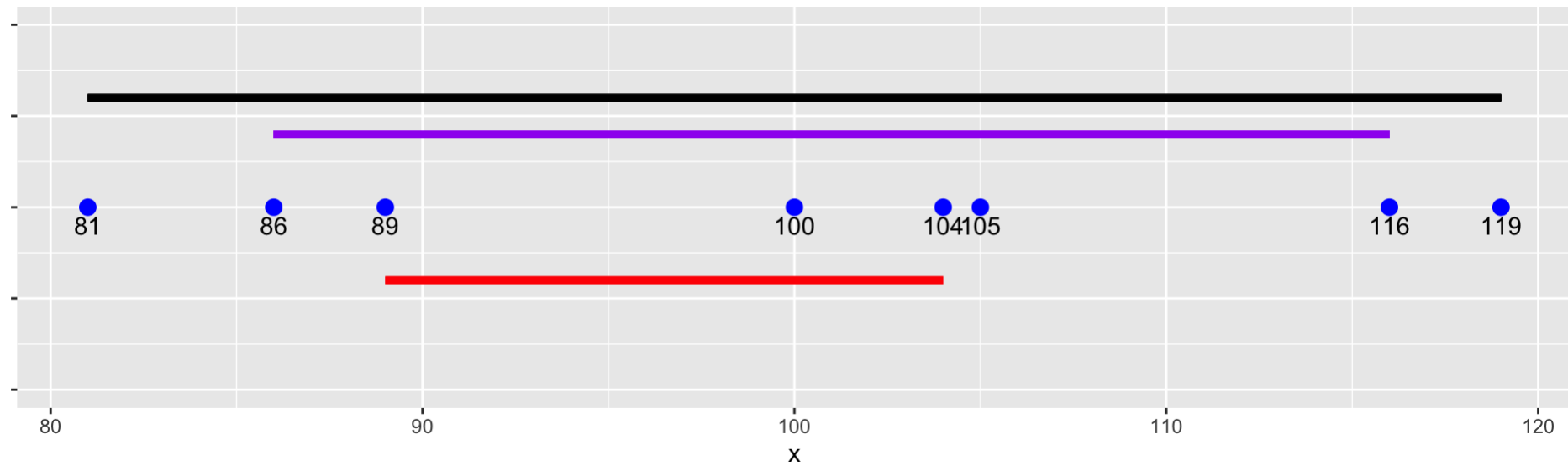

課堂練習 17

求新生兒體重的變異數與標準差。

2.03, 2.21, 2.29, 2.43, 2.44, 2.51, 2.54, 2.54, 2.55, 2.61, 2.62, 2.63, 2.64, 2.64, 2.66,
2.69, 2.70, 2.71, 2.72, 2.80, 2.80, 2.82, 2.87, 2.88, 2.88, 2.90, 2.92, 2.93, 2.96, 3.02,
3.02, 3.03, 3.04, 3.04, 3.05, 3.08, 3.12, 3.15, 3.16, 3.16, 3.17, 3.21, 3.22, 3.23, 3.24,
3.27, 3.30, 3.33, 3.37, 3.38, 3.40, 3.45, 3.46, 3.46, 3.49, 3.53, 3.55, 3.56, 3.61, 3.68,
3.70, 3.71, 3.72, 3.88, 4.20

方法二

位置參數間的距離代表全體資料的差異程度。



全距、四分位距和四分位差

- 全距=最大-最小
- 四分位距 $IQR = Q_3 - Q_1$
- 四分位差 $QD = \frac{Q_3 - Q_1}{2}$

全距、四分位距和四分位差

```
## [1] 12 13 14 14 15 15 15 15 16 16 16 16 16 16 17 17 18 18 18 19
```

全距、四分位距和四分位差

[1] 7

75%

2

75%

1

課堂練習 18

求新生兒體重的全距、四分位距和四分位差。

2.03, 2.21, 2.29, 2.43, 2.44, 2.51, 2.54, 2.54, 2.55, 2.61, 2.62, 2.63, 2.64, 2.64, 2.66,
2.69, 2.70, 2.71, 2.72, 2.80, 2.80, 2.82, 2.87, 2.88, 2.88, 2.90, 2.92, 2.93, 2.96, 3.02,
3.02, 3.03, 3.04, 3.04, 3.05, 3.08, 3.12, 3.15, 3.16, 3.16, 3.17, 3.21, 3.22, 3.23, 3.24,
3.27, 3.30, 3.33, 3.37, 3.38, 3.40, 3.45, 3.46, 3.46, 3.49, 3.53, 3.55, 3.56, 3.61, 3.68,
3.70, 3.71, 3.72, 3.88, 4.20