

Lecture 3 Summary (Chapter 2)

2) Measures of Center

Data set: x_1, x_2, \dots, x_n

(A) **Mode:** The number that occurs most often

(B) **Mean (Average)**

Sample mean: $\bar{x} = (x_1 + x_2 + \dots + x_n)/n = \sum_{i=1}^n x_i/n$

Here, $\sum_{i=1}^n = x_1 + x_2 + \dots + x_n$

Example 5 The birth weights in pounds of five babies born in a hospital on a certain day:

9.2 6.4 10.5 8.1 7.8

$$\bar{x} = \frac{9.2 + 6.4 + 10.5 + 8.1 + 7.8}{5} = \frac{42.0}{5} = 8.4 \text{ (pounds)}$$

(C) **Median (Middle value)**

Sample median: Middlemost value of the ordered data set

Algorithm for obtaining median

1. Order data

$$x_1, x_2, \dots, x_n \longrightarrow x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

Example 5 (continued)

$$x_1 = 9.2, \quad x_2 = 6.4, \quad x_3 = 10.5, \quad x_4 = 8.1, \quad x_5 = 7.8$$

$$\rightarrow x_{(1)} = 6.4, \quad x_{(2)} = 7.8, \quad x_{(3)} = 8.1, \quad x_{(4)} = 9.2, \quad x_{(5)} = 10.5$$

2. i) If n is odd, take the middle value.

$$n = 5: \quad 6.4 \quad 7.8 \quad \underline{8.1} \quad 9.2 \quad 10.5$$

$$\text{Sample median: } \tilde{x} = x_{((n+1)/2)} = x_{((5+1)/2)} = x_{(3)} = 8.1$$

- ii) If n is even, take the average of the two middle values.

Example 6 The number of days the first six heart transplant patients at Stanford survived after their operations:

$$15 \quad 3 \quad 46 \quad 623 \quad 126 \quad 64 \longrightarrow 3 \quad 15 \quad \underline{46} \quad \underline{64} \quad 126 \quad 623$$

$$\frac{46 + 64}{2} = \frac{110}{2} = 55$$

Sample median:

$$\begin{aligned} \tilde{x} &= (x_{(n/2)} + x_{((n/2)+1)})/2 = (x_{(6/2)} + x_{((6/2)+1)})/2 = (x_{(3)} + x_{(4)})/2 \\ &= (46 + 64)/2 = 55 \end{aligned}$$

$$\text{Sample mean: } (15 + 3 + 46 + 623 + 126 + 64)/6 = 146 \text{ (days)}$$

Only one patient lived longer than 146.2 days.

Median is a better indicator in this case.

- Sample mean: greatly influenced by outlying values
 \bar{x} : appropriate for a bell-shaped population
- Sample median: not greatly influenced by outlying values
 \tilde{x} : appropriate for income or salary data

Distribution

- Skewed to the left (left skewed): Long left tail, mean < median
- Skewed to the right (right skewed): Long right tail, mean > median
- Symmetric: mean = median
- Unimodal: one mode
- Bimodal: two modes
- Multimodal: multiple modes