

Introduction to Statistics

Haneul Lee, PT, DSc

What are statistics?

- ▶ They are simply tools we use to describe, organize, and interpret information (data)
- ▶ We collect the data , use **descriptive statistics** to describe it, and then **inferential statistics** to interpret it.



A few simple steps

1. Identify the problem by asking a question
2. State your hypothesis
3. Identify the independent variable (IV)
4. Identify the dependent variable (DV)
5. choose the right statistical test



Statistics – A new language!

- ▶ In stats we use words AND symbols to represent information
- ▶ We use statistics to help us draw conclusions about data



Key Terms

Measures of central tendency

- ▶ Mean ***M or \bar{x}***
- ▶ Median ***Md***
- ▶ Mode ***Mo***



Mean (M / \bar{x})

- ▶ Arithmetic mean
- ▶ Most common type of average
- ▶ Calculated by adding all of the values up and then dividing the total by the number of values



Mean (M / \bar{x})

Mean for the sample


$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Mean for the Population

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i$$

Mean (M / \bar{x})

Locations	Number of COPD patients
Loma Linda University Medical Center	1065
San Bernardino Medical Center	785
Redlands Community Hospital	453
Riverside County Regional Medical Center	985

- To calculate the mean, you add 1065 plus 785 plus 453 plus 985 and divided the sum by 4 (n=4)
 - The mean (average) is 822
-
- 

Median (Md)

- ▶ Middle value of ordered data
- ▶ If the number of values is **odd**, the median is the number located in the exact middle of the ordered list
- ▶ If the number of values is **even**, the median is found by computing the mean of the two middle numbers



Median (Md)

Car	Cost
Ford Mustang GT	\$ 38,789.00
Audi A4	\$ 41,266.00
GMC sierra Denali	\$ 42,537.00
BMW 335i	\$ 46,289.00

Car	Cost
Ford Mustang GT	\$ 38,789.00
Audi A4	\$ 41,266.00
GMC sierra Denali	\$ 42,537.00
Nissan 370Z Nismo	\$ 43,628.00
BMW 335i	\$ 46,289.00

- The median for the first data set is \$ 42,447.00.
$$= (41,266 + 42,537) / 2$$
 - The median for the second data set is \$ 42,537.00.
-



Mean Vs. Median

- ▶ Why use the median as a measure of central tendency?

Median is not sensitive to *extreme scores* (outliers)

*** Don't use mean when there are outliers**



Mean Vs. Median

Car	Cost
Ford Mustang GT	\$ 38,789.00
Audi A4	\$ 41,266.00
GMC sierra Denali	\$ 42,537.00
Nissan 370Z Nismo	\$ 43,628.00
Nissan GT-R	\$ 84,060.00

- The mean for this data set is \$
 - The median is \$
-



Mode (M_o)

- ▶ Most frequent value (there may be more than one mode)
- ▶ Mode is the best representation when
 - ✓ choosing the most popular response
 - ✓ Analyzing categories of data



Mode (M_o)

- ▶ Number of Children in the Family , $n=8$

0 1 4 1 6 1 2 3 2

Order : **0 1 1 1 2 2 3 4 6**



MODE



Mean

Average



Median

Middle



Mode

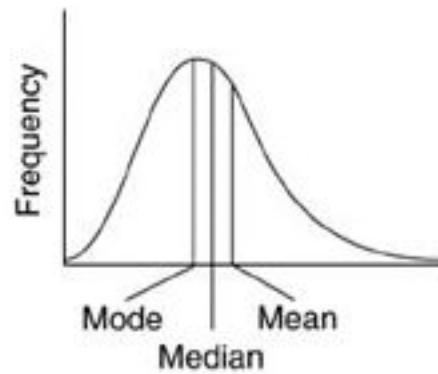
Most



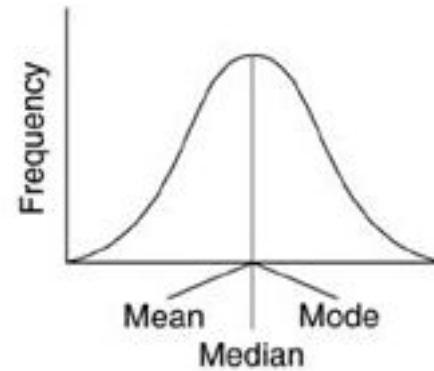
Cheat

Sheet!

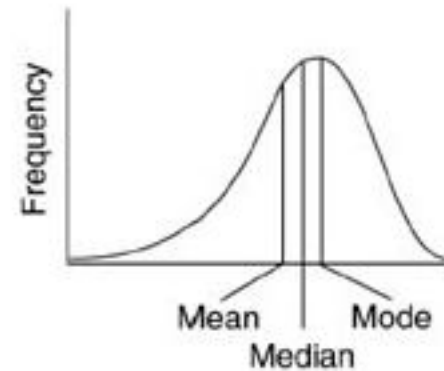
Symmetry and Skewness



c. Positively skewed distribution



a. Symmetrical distribution



b. Negatively skewed distribution



Symmetry and Skewness

Comparing Exam Scores

	Hard	Easy
4	89	9
5	35968	
6	225567789	67
7	34457799	246678
8	12	34456789
9	6	356678








Midterm Scores

41 47 52.5 53.5 58 58 58.5 59.5 61 70

- ▶ **Mean** 55.9
- ▶ **Median** 58
- ▶ **Mode** 58
- ▶ **Skewed?** Negatively skewed



Measures of Central Tendency and Levels of Measurement

Data	Central Tendency/Location		
	Mean	Median	Mode
Nominal			
Ordinal			
Quantitative			

Don't use if distribution is skewed

References

- ▶ Neil J. Salkind, *Statistics for People Who Have Statistics*, 4th edition, Sage
- ▶ Dianne V. Jewell, *Guided to Evidence-Based Physical Therapist Practice*, 3rd edition, Jones Bartlett Learning
- ▶ David S. Moore, *The basic Practice of Statistics*, 5th edition, Freeman
- ▶ Daniel T. Larose, *Discovering the Fundamentals of Statistics*, 2nd edition, Freeman
- ▶ Andy Field, *Discovering statistics using IBM SPSS Statistics*, 4th edition, Mobile study

