STA291

THURSDAY, 24 SEPTEMBER 2009

Announcement

- Exam 1: September 30th at 5pm to 7pm. Location MEH, Memorial Auditoriam. The make-up will be at 7:30pm to 9:30pm at the 8th floor of POT. You have to let me know if you want to take the make up by the midnight of Sept 27th via email.
- Calculator will be allowed to use in the exam but no open book nor open notes (no cell phone nor computer as well).

Measures of Central Location

- Also called Central *Tendency*
- "What is a typical measurement in the sample/population?"
  - Mean: Arithmetic average
  - Median: Midpoint of the observations when they are arranged in increasing order
  - Mode: Most frequent value

Mean and Median

- For skewed distributions, the median is often a more appropriate measure of central tendency than the mean
- The median usually better describes a “typical value” when the sample distribution is highly skewed

- Example:

  Monthly income for five persons (n = 5)

  1,000 2,000 3,000 4,000 100,000

- Median monthly income: 3000

Mean and Median

- Is there a compromise between the median and the mean? Yes!
- Trimmed mean:
  1. Order the data from smallest to largest
  2. Delete a selected number of values from each end of the ordered list
  3. Find the mean of the remaining values
- The trimming percentage is the percentage of values that have been deleted from each end of the ordered list.

Mode

- Mode of a data set is the most frequently occurring value
- Can speak of a data set being *unimodal*, *bimodal*, or *multimodal*
- Can be calculated on nominal (!) data
- On a histogram, where would the mode be?

Summary: Measures of Location

Mean- Arithmetic Average

- Mean of a Sample - $\bar{x}$
- Mean of a Population - μ

Can be calculated only on quantitative data

Notation: Subscripted variables
n = # of units in the sample
N = # of units in the population
x = Variable to be measured
x_i = Measurement of the *ith* unit

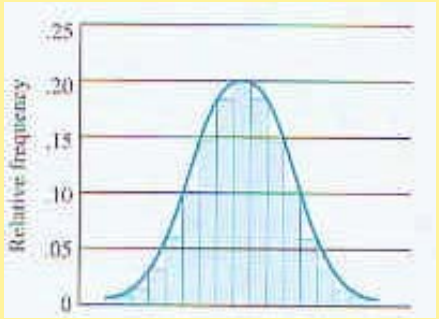
Median – Midpoint of the observations when they are arranged in increasing order

Can be calculated on *quantitative* **or** *ordinal* data

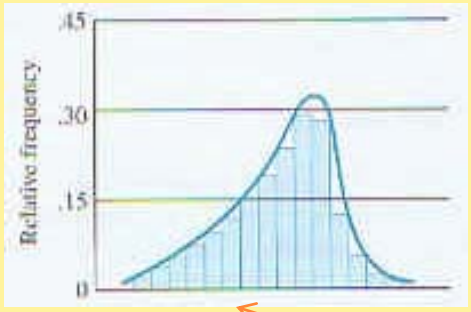
Mode- Most frequent value.

Can be calculated on quantitative, ordinal, or nominal data!

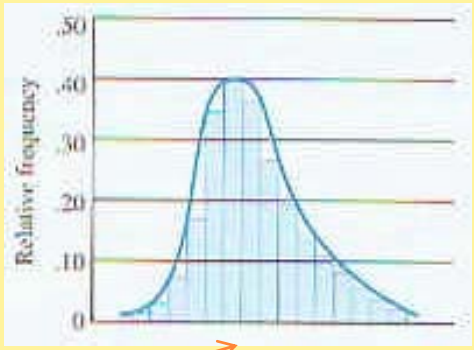
Review: Shapes of Distributions



Symmetric Distribution



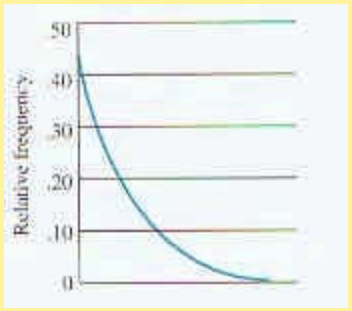
Skewed to the left



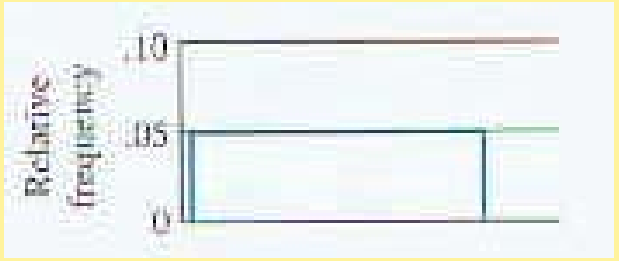
Skewed to the right

Examples

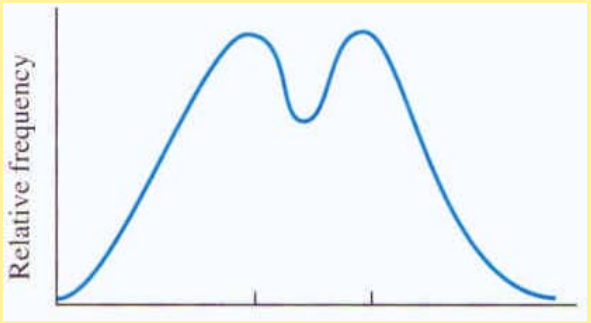
Exponential Distribution



Uniform Distribution



Bimodal Distribution



Summarizing Data Numerically

- Center of the data
  - Mean
  - Median
  - Mode
- Dispersion of the data
  - Variance, Standard deviation
  - Interquartile range
  - Range

Mean vs. Median vs. Mode

- The mean is sensitive to outliers, median and mode are not
- In general, the median is more appropriate for skewed data than the mean
- In some situations, the median may be too insensitive to changes in the data
- The mode may not be unique

Mean vs. Median vs. Mode

- Mean: Interval data with an approximately symmetric distribution

- Median: Interval or ordinal data

- Mode: All types of data

Mean and Median

- Example: For towns with population size 2500 to 4599 in the U.S. Northeast in 1994, the mean salary of chiefs of police was $37,527, and the median was $30,500.

- Does this suggest that the distribution of salary was skewed to the left, symmetric, or skewed to the right?

Mean, Median, Mode—Another Example

Response	Frequency	Relative Frequency
every day	969	
a few times a week	452	
once a week	261	
less than once a week	196	
Never	76	
TOTAL		

- Identify the mode
- Identify the median response
- Mean?

Attendance Survey Question #9

- On an index card
  - Please write down your name and section number
  - Today's Questions: