# Biostatistics

**Lecture 3**
**Prepared by Dr. Ibrahim AL-Jaafari**

# Organizing and Displaying Data Quantitative Data

- **Frequency Distributions**

- **Constructing Frequency Distribution Tables**

- **Relative and Percentage Distributions**

- **Graphing Grouped Data**

  - **Histograms**

  - **Polygons**

# Frequency Distributions

**Definition**

A *frequency distribution* for quantitative data lists all the classes and the number of values that belong to each class. Data presented in the form of a frequency distribution are called *grouped data*.

# Class Boundary

**Definition**

The *class boundary* is given by the midpoint of the upper limit of one class and the lower limit of the next class.

# Class Boundary

| Age | Frequency | Class Boundaries |
|---|---|---|
| 0 – 9 | 20 | –0.5 to 9.5 |
| 10 – 19 | 15 | 9.5 to 19.5 |
| 20 – 29 | 25 | 19.5 to 29.5 |
| 30 – 39 | 20 | 29.5 to 39.5 |
| 40 – 49 | 20 | 39.5 to 49.5 |
| 50 – 59 | 30 | 49.5 to 59.5 |
| 60 – 69 | 20 | 59.5 to 69.5 |

- Class boundary for first class (0-9)

= 9 + 10 / 2 = 9.5

- Class boundary for first class (10-19)

= 19 + 20 / 2 = 19.5

- Class boundary for first class (20-29)

= 29 + 30 / 2 = 29.5

# Class Width – Class Size

- **The difference between the two boundaries of the class**

**Class width = Upper Boundary – Lower Boundary**

# Class Width – Class Size

| Age | Frequency | Class Boundaries | Class Width |
|---|---|---|---|
| 0 - 9 | 20 | -0.5 to 9.5 | 10 |
| 10 - 19 | 15 | 9.5 to 19.5 | 10 |
| 20 - 29 | 25 | 19.5 to 29.5 | 10 |
| 30 - 39 | 20 | 29.5 to 39.5 | 10 |
| 40 - 49 | 20 | 39.5 to 49.5 | 10 |
| 50 - 59 | 30 | 49.5 to 59.5 | 10 |
| 60 - 69 | 20 | 59.5 to 69.5 | 10 |

Class width for the first class (0-9)

$= 9.5 - 0.5 = 9$

$= 19.5 - 10.5 = 9$

$= 29.5 - 20.5 = 9$

$= 39.5 - 30.5 = 9$

$= 49.5 - 40.5 = 9$

# Calculation of Class Width

- **Approximate Class Width** $= \dfrac{Largest\ value - Smalest\ value}{Number\ of\ Classes}$

# Example for Calculation of Class Width

**Construct a frequency distribution table for these data**

8, 25, 11, 15, 29, 22, 10, 5, 17, 21, 22, 13, 26, 16, 18, 12, 9, 26, 20, 16, 23, 14, 19, 23, 20, 16, 27, 16, 21, 14

# Solution

**Maximum value = 29**                    **Minimum value = 5**

Suppose we decide to group these data using five classes of equal width

**Approximate Class Width = 29 – 5 / 5 = 4.8**

**5–9, 10–14, 15–19, 20–24, and 25–29**

# Class Width Table

| |
| --- |
| 5 - 9 |
| 10 - 14 |
| 15 – 19 |
| 20 – 24 |
| 25 - 29 |

# Class Midpoint – Mark

**Calculating Class Midpoint or Mark**

$$\text{Class Midpoint or Mark} = \frac{Lower\ limit + Upper\ Limit}{2}$$

# Class Midpoint Class Mark

| Age | Frequency | Class Boundaries | Class Width | Class Midpoint |
|---|---|---|---|---|
| 0 - 9 | 20 | -0.5 to 9.5 | 10 | $0 + 9 \div 2 = 4.5$ |
| 10 - 19 | 15 | 9.5 to 19.5 | 10 | $10 + 19 \div 2 = 14.5$ |
| 20 - 29 | 25 | 19.5 to 29.5 | 10 | $49 \div 2 = 24.5$ |
| 30 - 39 | 20 | 29.5 to 39.5 | 10 | $69 \div 2 = 34.5$ |
| 40 - 49 | 20 | 39.5 to 49.5 | 10 | $89 \div 2 = 44.5$ |
| 50 - 59 | 30 | 49.5 to 59.5 | 10 | $109 \div 2 = 54.5$ |
| 60 - 69 | 20 | 59.5 to 69.5 | 10 | $129 \div 2 = 64.5$ |

- **Class Midpoint**
- $0 + 9 / 2 = 4.5$
- $10 + 19 / 2 = 14.5$
- $49 / 2 = 24.5$
- $69 / 2 = 34.5$

# Relative Frequency and Percentage Distributions

**What is cumulative frequency distribution ??**

The total number of values that fall below the upper boundary of each class

**What is cumulative relative frequency ??**

Cumulative frequency ÷ Total observations in the data set

**What is cumulative percentage ??**

Cumulative relative frequency × 100

# Relative Frequency and Percentage Distributions

| Age | Number of Patients | Cumulative Frequency | Cumulative Relative Frequency | Cumulative Percentage |
|---|---|---|---|---|
| 0 - 9 | 20 | 20 | 20÷130 = 0.153 | 15.3 % |
| 10 - 19 | 15 | 20+15 =35 | 35÷130 = 0.269 | 26.9 % |
| 20 - 29 | 25 | 20+15+25=60 | 60÷130 = 0.461 | 46.1 % |
| 30 - 39 | 20 | 20+15+25+20=80 | 80÷130 = 0.615 | 61.5 % |
| 40 - 49 | 20 | 20+15+25+20+20=100 | 100÷130 = 0.769 | 76.9 % |
| 50 - 59 | 30 | 20+15+25+20+20+30= 130 | 130÷130 = 1 | 100 |

# GRAPHING QUANTITATIVE DATA

**How to present quantitative data ??**

**1- Histograms /** a graph in which classes are marked on horizontal axis and the frequencies are marked on the vertical axis which represent the height of bars.

In a histogram, the bars are attached to each other.

# Histograms
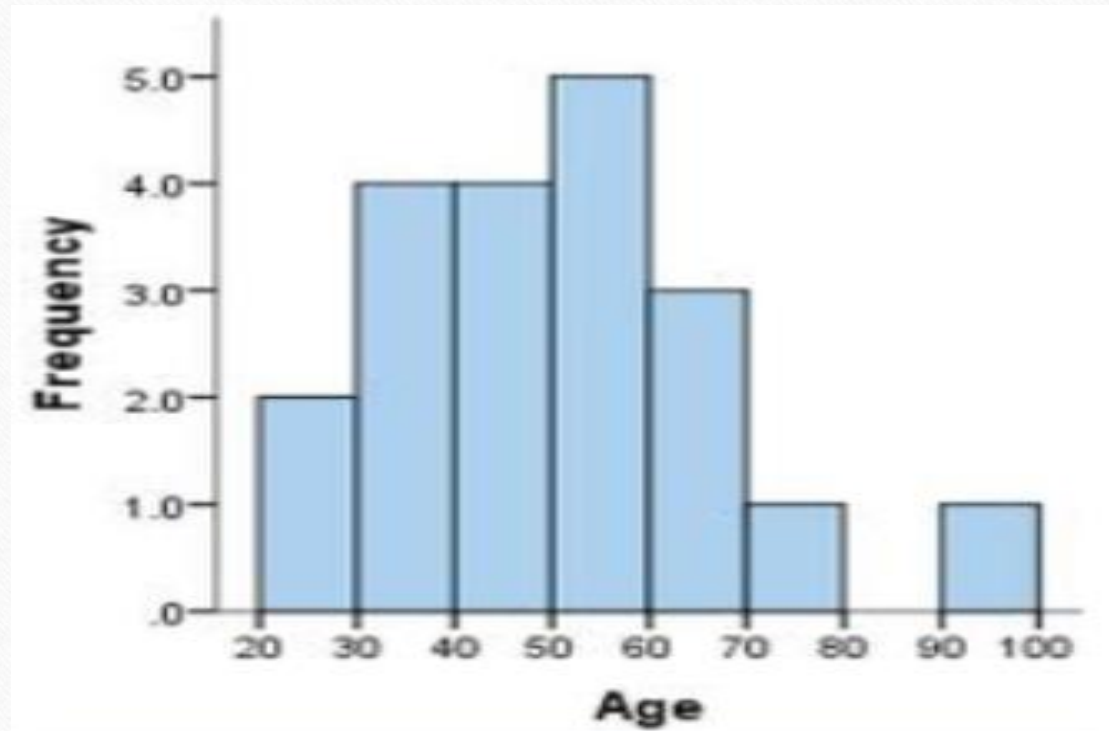
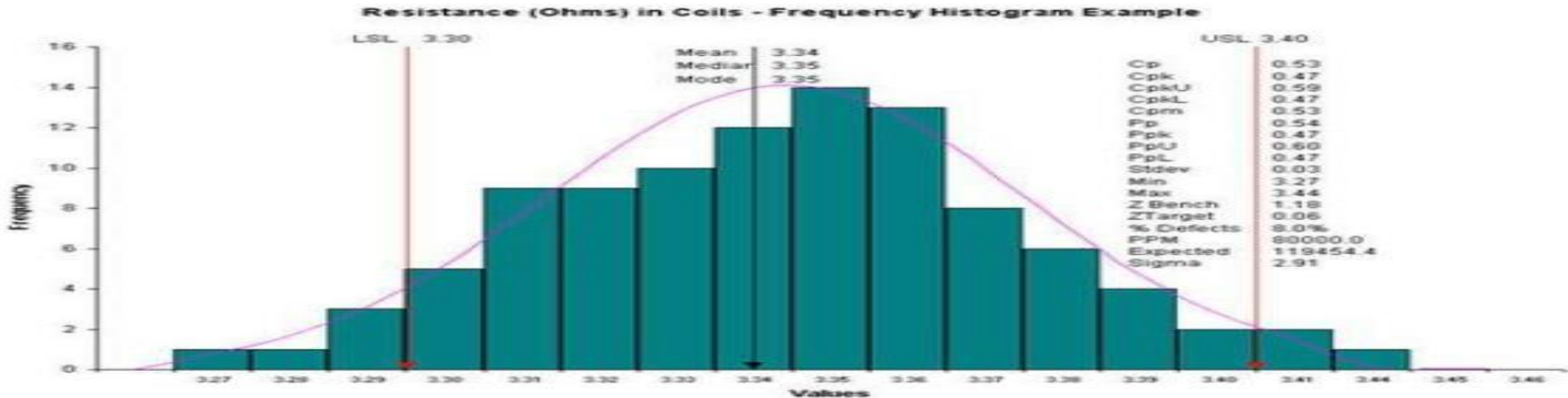# Histograms

**What are the shapes of histograms ??**

**A-) Symmetric histogram** / is identical on both sides of its central points.



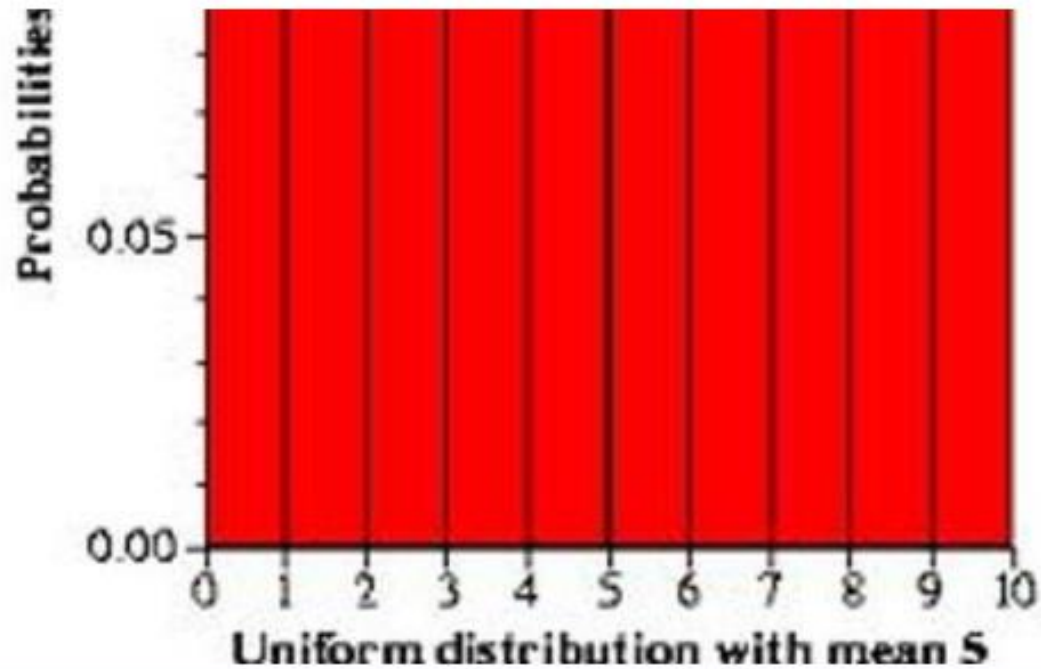Resistance (Ohms) in Coils - Frequency Histogram Example

# Histograms



**B-) Bimodal symmetric histogram / is identical on both sides of its central points with two modes**
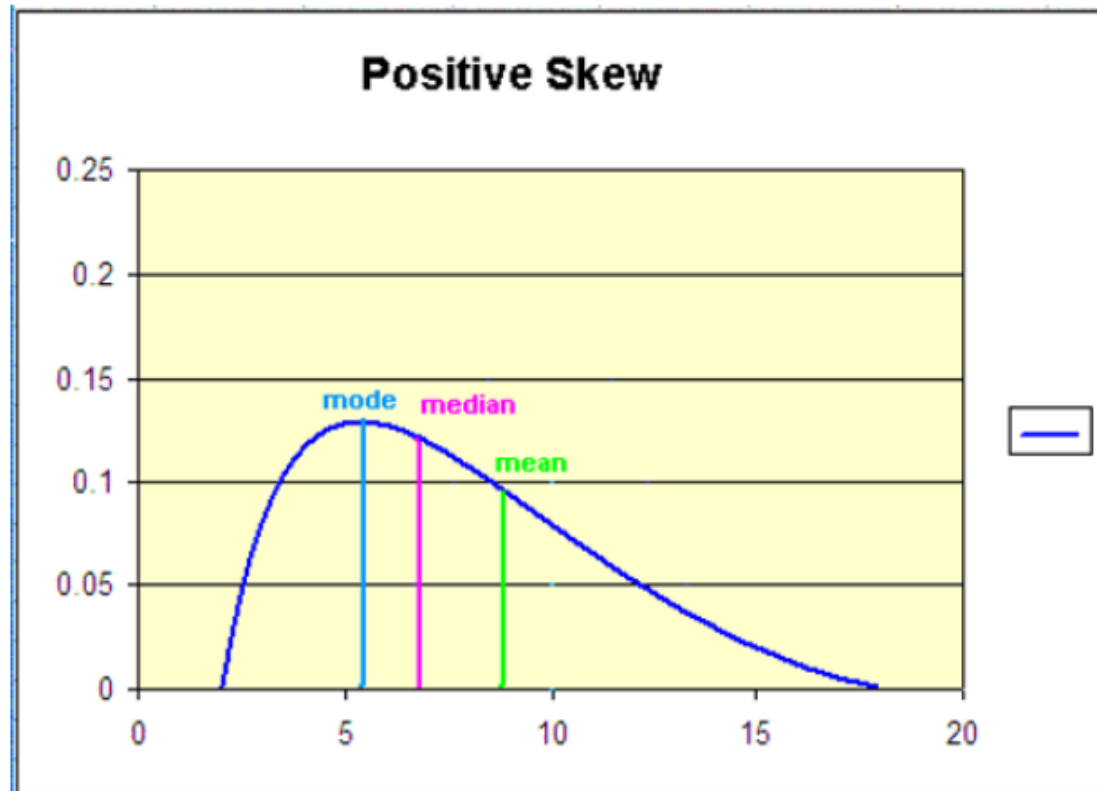
# Histograms

**C-) Uniform or rectangular histogram ( symmetric ) /** the frequencies of each class are the same or equal to each other.
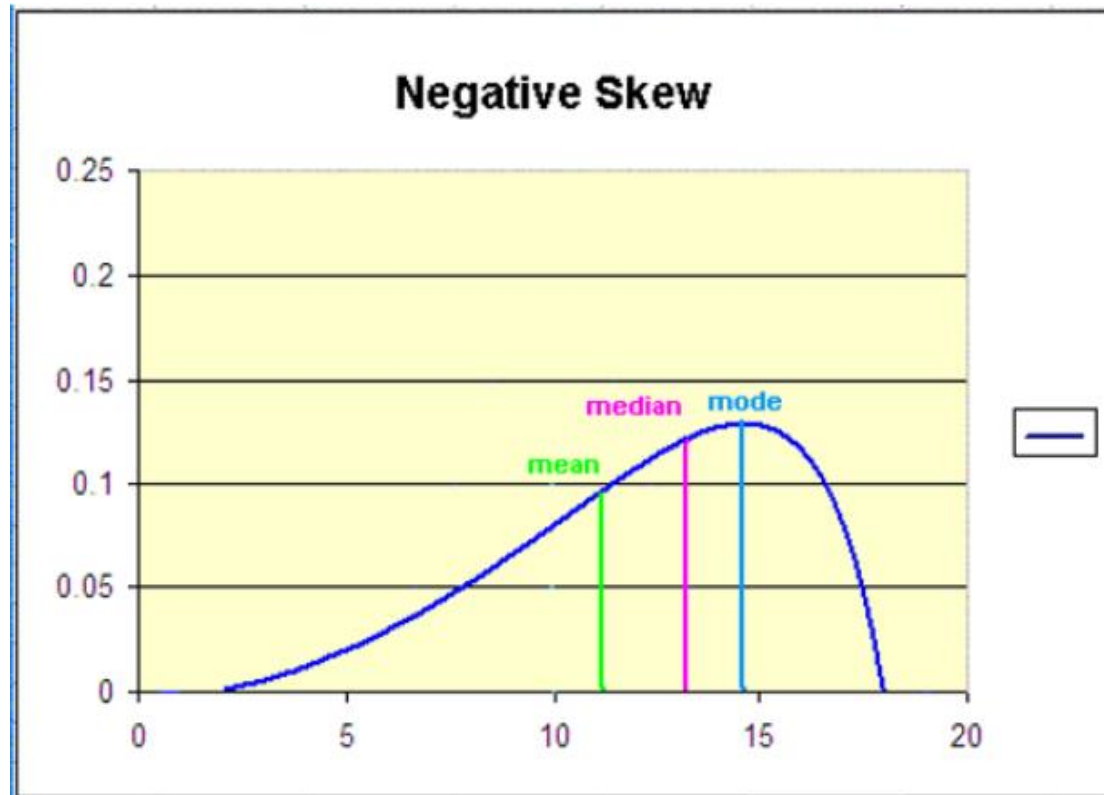
# Histograms



**D-) Skewed to the right histogram ( positive skewed ) /** Most of data is shown in the left side of histogram and the tail on the other right side.

# Histograms



Negative Skew

E-) Skewed to the left histogram ( negative skewed ) / Most of data is shown in the right side of histogram and the tail on the other left side.

# Polygons

**2- Polygons /** A graph formed by joining the midpoints of the tops of bars in a histogram with straight lines.

# Polygons

# Stem and Leaf display

**3- Stem and Leaf display /** each value is divided into two portions -- a stem and a leaf. Then the leaves for each stem are shown separately in a display.

# Stem and Leaf display

**Construct a stem-and-leaf display for these data**

Example / 22 , 26 , 27 , 31 , 33 , 35 , 42 , 44 , 46 , 57 , 58 , 59 , 61 , 63, 64 , 65 ,67

# Solution

| Stem | Leaf |
|------|------|
| 2 | 2 6 7 |
| 3 | 1 3 5 |
| 4 | 2 4 6 |
| 5 | 7 8 9 |
| 6 | 1 3 4 5 7 |

# Reference

- **Prem S. Mann 1998, Introductory Statistics, 7<sup>th</sup> edn, New York, USA.**

# Good Luck for All Students

- Please do not hesitate to contact me if you have any questions.

- Dr. Ibrahim AL-Jaafari

- www.Alghamdi-Biostatistics.com

- Email. Bio-stat@Hotmail.com

- Mobile Number : 0553777925

# سبحان الله وبحمده سبحان الله العظيم

ذكر الله أعظم ما في الوجود ،، لعل الله يرحمنا بعلم تعلمناه في الحياة الدنيا

أستغفر الله