

## SRF Elementary Statistics Research

### Frequency Diagram

-An organization of raw data in the form of a table using classes and frequencies.

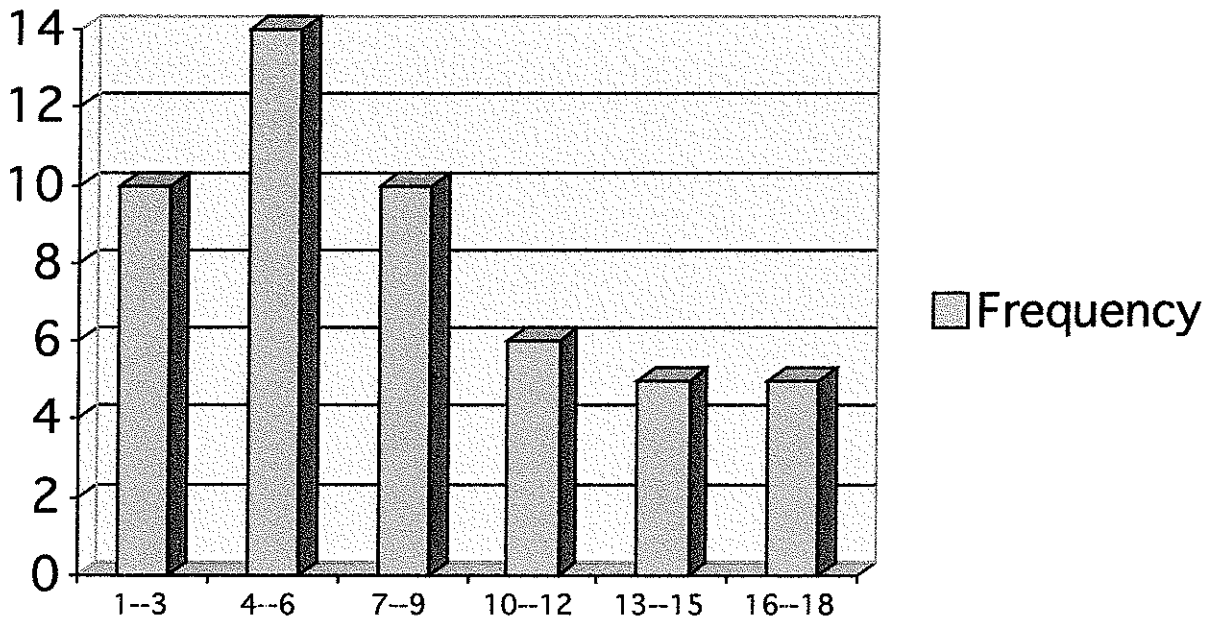
Example:

Class Limits (miles to traveled to work)	Frequency
1-3	10
4-6	14
7-9	10
10-12	6
13-15	5
16-18	5

### Histogram

-A representation of a frequency distribution by means of rectangles whose widths represent class intervals and whose heights represent corresponding frequencies.

Example:



### Distribution

-An arrangement of values of a variable showing their observed or theoretical frequency of occurrence

### Mean

-The sum of a group of numbers divided by the number of numbers; an average.

### Median

-The middle value in a distribution, above and below which lie an equal number of values.

### Mode

-The value or item occurring most frequently in a series of observations or statistical data.

### Standard Deviation

-The standard deviation is a measure of how spread out your data are.

The steps to calculate standard deviation are as follows:

1. Compute the mean for the data set.
2. Compute the deviation by subtracting the mean from each value.
3. Square each individual deviation.
4. Add up the squared deviations.
5. Divide by one less than the sample size.
6. Take the square root.

Formula for Standard Deviation:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

### Kurtosis

-Is a measure of *peakedness* in a distribution, that is, how “high” the peak is in relation to the rest of the data.

-Kurtosis “excess” is simply the kurtosis minus three and is denoted  $y_2$  or  $b_2$ . This form is most commonly used because the  $y_2$  of a normal distribution is zero. So this sort of serves as a standardization, from which, all other data can be compared.

-Keeping in mind that the  $y_2$  of a normal distribution is zero, we can infer that A distribution with a high peak ( $y_2 > 0$ ) is called leptokurtic, a flat-topped curve ( $y_2 < 0$ ) is called platykurtic, and the normal distribution ( $y_2 = 0$ ) is called mesokurtic.

### Calculations for Kurtosis Excess for several distributions

### Skewness

-Is a measure of the degree of asymmetry of a distribution, that is to say, it how well the right side of the data “matches” the left side of the data.

-If the left tail is more pronounced than the right tail, the function is said to have negative skewness. If the opposite is true, it has positive skewness. If the two sides are equal, it has zero skewness.

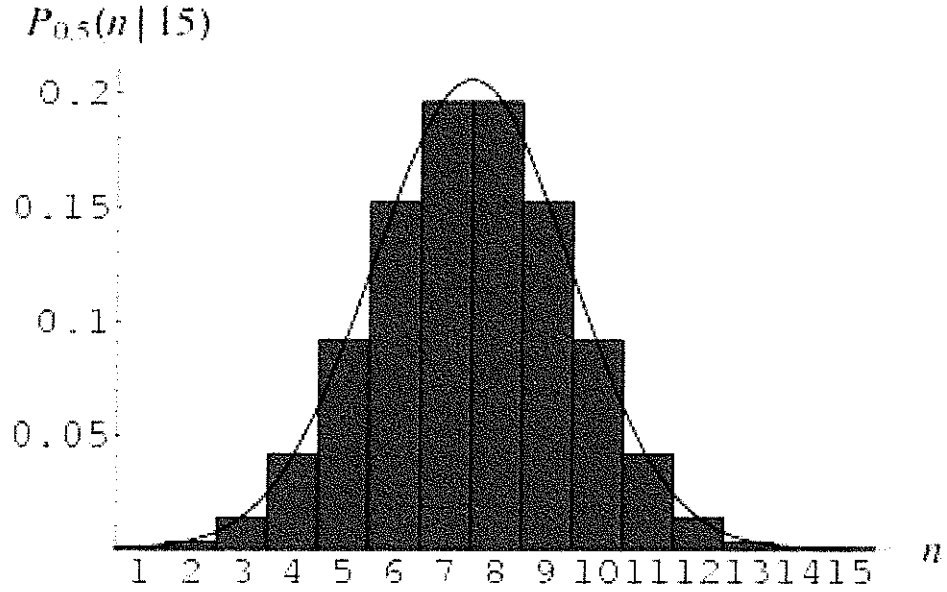
-There are a number of ways to measure symmetry in a distribution.

### Calculations for skewness in several distributions

Symmetric Distribution

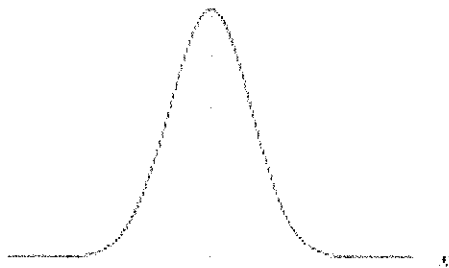
-Is defined as a distribution in which the data values are uniformly distributed about the mean.

-There are an infinite number of symmetric distributions, a couple can be seen below



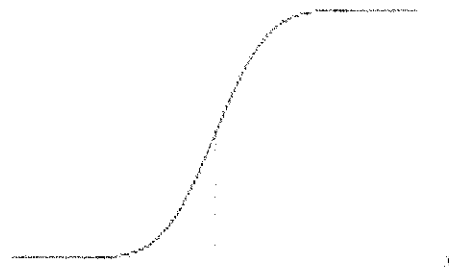
Symmetric

$P(x)$



Not Symmetric

$D(x)$



## Tailedness

-A description of the "tails" of data.

Different Tails on Data

