Correlation

Dr. Sabiha Khan

Correlation and Causation

- The following factors should be examined to interpret the nature and extent of relationship between two or more variables:
- Chance coincidence:
- Influence of third variable
- Mutual influence

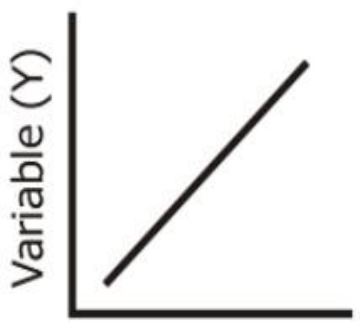
Types of Correlation

1. Positive and Negative Correlation

- A **positive correlation** is a relationship between two variables in which both variables move in the same direction. Therefore, when one variable increases as the other variable increases, or one variable decreases while the other decreases. An example of positive correlation would be height and weight. Taller people tend to be heavier.
- A **negative correlation** is a relationship between two variables in which an increase in one variable is associated with a decrease in the other. An example of negative correlation would be height above sea level and temperature. As you climb the mountain (increase in height) it gets colder (decrease in temperature).

- **Linear Correlation**
- There exists a linear correlation if the ratio of change in the two variables is constant.
- If we plot these coordinates on a graph, we will get a straight line.

X	:	Y
1	:	10
2	:	20
3	:	30



- **Non- Linear or Curvilinear Correlation**
- There exists a curvilinear correlation if the change in the variables is not constant.
- If we plot these coordinates on a graph, we will get a curve.

X	:	Y
1	:	10
2	:	15
3	:	40

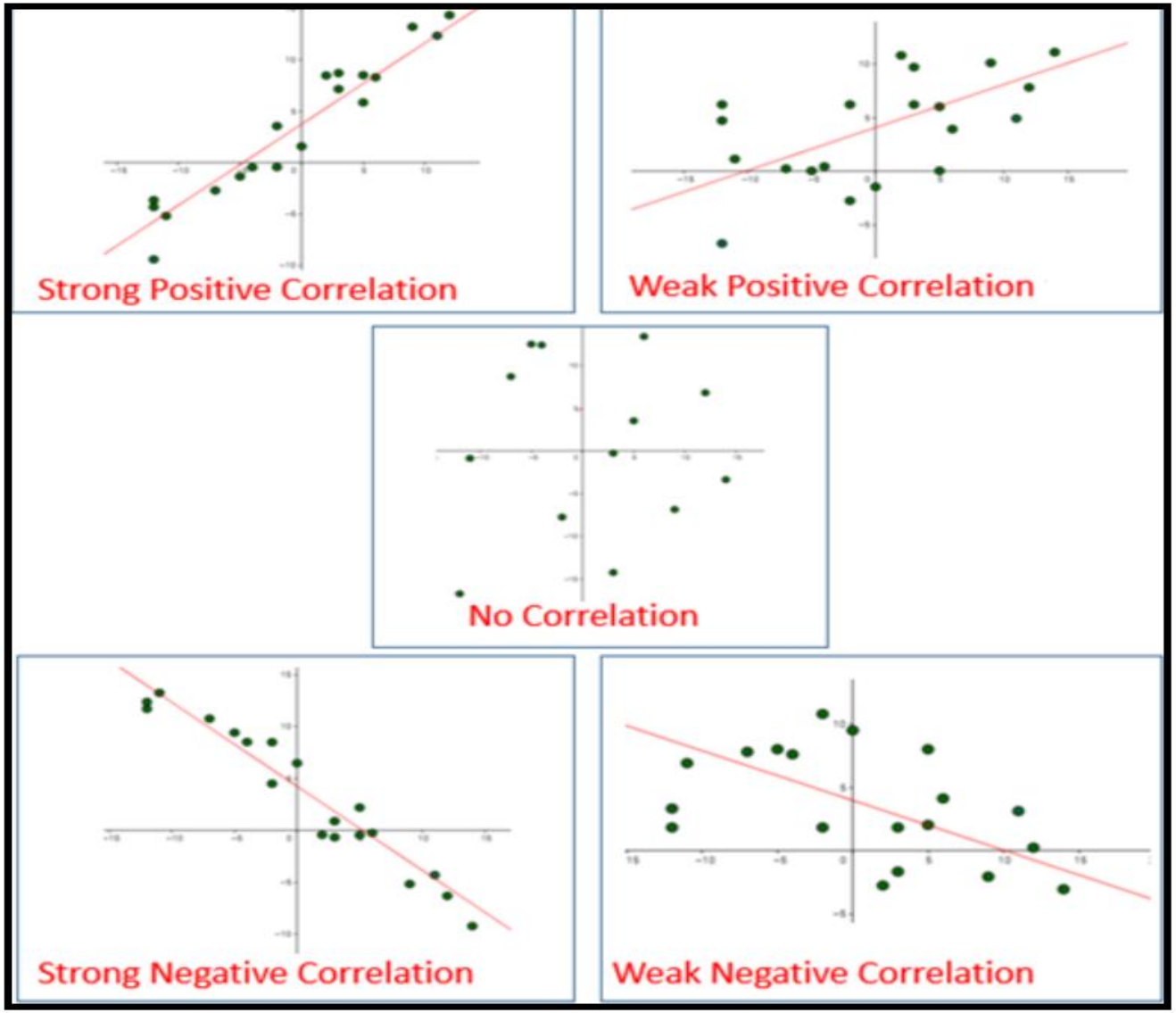
3. Simple, Partial and Multiple

- **Simple correlation:** When we consider only two variables and check the correlation between them, it is termed as simple correlation. For example, the radius and circumference of a circle.
- **Multiple correlation:** When we consider three or more variables for correlation, it is termed as multiple correlation. For example, the price of cola drink, temperature, income, and the demand for cola.
- **Partial correlation:** When one or more variables are kept constant and the relationship is studied between the others, it is termed as partial correlation. For example, if we keep the price of cola constant and check the correlation between temperature and the demand for cola, it is a partial correlation.

Methods of Correlation Analysis

- 1. Scatter Diagram Method
- 2. Karl Pearson ’s Coefficient of Correlation Method
- 3. Spearman’s Rank Correlation Method
- 4. Method of Least Squares

Scatter Diagram Method



Types of Correlation Coefficients:

Coefficient	Conditions applied for use
Ø (phi)	Both x & y variables are measured on a nominal scale.
P (rho)	Both x & y variables are measured on, or changed to, a ordinal scale.
r	Both x & y variables are measured on an interval or ratio scale.

Karl Pearson - Correlation co-efficient

Age (x)	43	21	25	42	57	59
Glucose level (Y)	99	65	79	75	87	81

Solution

x	$d_x = x - \bar{x}$	d_x^2	y	$d_y = y - \bar{y}$	d_y^2	$d_x \times d_y$
43	1.83	3.35	99	18	324	32.94
21	-20.17	406.83	65	-16	256	322.72
25	-16.17	261.47	79	-2	4	32.34
42	0.83	0.69	75	-6	36	-4.98
57	15.83	250.59	87	6	36	94.98
59	17.83	317.91	81	0	0	0
Σx = 247		Σd_x^2 = 1240.87	486		Σd_y^2 = 656	$\Sigma d_x d_y$ = 478

$$\bar{x} = \frac{\Sigma x}{N} = \frac{247}{6} = 41.17$$

$$\bar{y} = \frac{\Sigma y}{N} = \frac{486}{6} = 81$$

$$r = \frac{\Sigma d_x d_y}{\sqrt{\Sigma d_x^2 \times \Sigma d_y^2}}$$

$$r = \frac{478}{\sqrt{1240.87 \times \sqrt{656}}}$$

$$= \frac{478}{85.23 \times 25.61} = \frac{478}{902.24} = 0.53$$

positive

Spearman - correlation co-efficient.

IQ (x)	106	100	86	101	99	103	97	113	112	110
Hrs. of T.V. per week (y)	7	27	2	50	28	29	20	12	6	17

Solution

x	y	Rank x_i	Rank y_i	Rank difference d_i	d_i^2
86	2	1	1	0	0
97	20	2	6	-4	16
99	28	3	8	-5	25
100	27	4	7	-3	9
101	50	5	10	-5	25
103	29	6	9	-3	9
106	7	7	3	4	16
110	17	8	5	3	9
112	6	9	2	7	49
113	12	10	4	6	36
					$\Sigma d_i^2 = 194$

$$\rho = 1 - \frac{6(\Sigma d_i^2)}{n(n^2-1)}$$

$$= 1 - \frac{6(194)}{10(10^2-1)} = 1 - \frac{6 \times 194}{1000-10}$$

$$= 1 - \frac{1164}{990} = 1 - 1.175 = \underline{-0.175}$$

very low correlation

THANK YOU

- Disclaimer: The content displayed in the PPT has been taken from variety of different websites and book sources. This study material has been created for the academic benefits of the students alone and I do not seek any personal advantage out of it.