Name: Omar faruk

Registration Number: 2019831055

STA - 101W

Part: A

---

Ans: to the que: NO - 1

(i)

By population we mean total of units/items/individuals which are under investigation according to some pre-determined objective and are available in a specified area at a specific time period. If the objective is to study the socioeconomic characteristic of patients visiting out door of a hospital, then all the patients population.

(ii)

A representative number of population units which are under investigation constitutes a sample. For example,

if out door patients of a hospital in a month constitute a population, then patients visited in some selected days ~~nurses working in a particular~~ constitute a sample. If nurses working in a hospital constitute a population, then nurses working in a particular working hours of the day may be a sample.

(iii)

An the information of a particular characteristic vary from unit to unit, is called variable. variable is of two types, ⓐ Qualitative Variable, ⓑ Quantitative Variable.

(iv)

Data which are collected from sampling units by personal interview or by mailed questionnaire are known as primary data. These data, if

are not recorded or stored in a systematic fashion, are called 'Raw Data'.

(v)

A frequency distribution is a list, table or graph that displays the frequency of various outcomes in a sample. Each entry in the table contains the frequency or count of the occurrences of values within a particular group or interval.

(vi)

Graphical representation of data refers to the use of charts and graphs to visually display, analyze, clarify and interpret numerical data, functions and other quantative structures. The graphical representation of data and information offers a quick and simple way of understanding the features and drawing comparisons.

## Ans: to the que: No-2

### (a)

**Measures of dispersion:**

Measures of dispersion indicates the scattering of data, explains the disparity of data from one another, delivering a precise view of their distribution. The measures of dispersion display and gives us an idea about the variation and the central value of an individual item. It is also defined as an average deviation of observation from some central value.

To understand the consistency, homogenity we need to study the measures of dispersion. We also need measures of dispersion to control the quality of product in a production house. It is important to describe the spread of the data or its variation around a central value.

The main importance of measures of dispersion is to study the variability of observations. The other important aspects of studying dispersion, are -

(i) Controlling variability in the experimental results.

(ii) Comparing the variability of two or more sets of observations.

(iii) Calculating the lower limit and upper limit of the observations under study.

## Ans: to the que: NO-2

(b)

Let, $x_1, x_2, x_3, \ldots\ldots, x_n$ be a set of $n$ observation.

Hence, $\sum x_i = \frac{n(n+1)}{2}$

Average, $\bar{x} = \frac{\sum x_i}{n} = \frac{n(n+1)}{2n} = \frac{n+1}{2}$

Variance, $\sigma^2 = \frac{1}{n}\sum(x_i - \bar{x})$

$= \frac{1}{n}\left[\sum(x_i^2 - 2x_i\bar{x} + \bar{x}^2)\right]$

$= \frac{1}{n}\left[\sum x_i^2 - 2\bar{x}\sum x_i + \sum\bar{x}^2)\right]$

$= \frac{1}{n}\left[\frac{n(n+1)(2n+1)}{6} - \frac{2(n+1)}{2} \times \frac{n(n+1)}{2} + n\times\left(\frac{n+1}{2}\right)^2\right]$

$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{2} + \frac{(n+1)^2}{4}$

$= \frac{n+1}{2}\left[\frac{2n+1}{3} - (n+1) + \frac{(n+1)}{2}\right]$

$= \frac{n+1}{2}\left[\frac{2n+1}{3} - \frac{n+1}{2}\right]$

$$= \frac{n+1}{2} \times \frac{4n+2-3n-3}{6}$$

$$= \frac{n+1}{2} \times \frac{n-1}{6}$$

$$= \frac{n^2-1}{12}$$

[showed]

## Ans: to the que: NO-3

(a)

Correlation describes the relationship between two or more variables are correlated if the change in one is followed by a change in the other. Correlation shows if the relationship is positive or negative and how strong the relationship is. Positive correlation shows describe the relationship between two variable which change together while a negative correlation describes the relation between two variables which change in opposite direction.

**Positive correlation:** When related variables move in the same direction, their relationship is positive. This correlation is measured by the coefficient of correlation "r". When r is greater than 0, it is positive. When r is 1, it is a perfect positive correlation. For example, the more money spent on expenditure will buy, the prices will be hike.

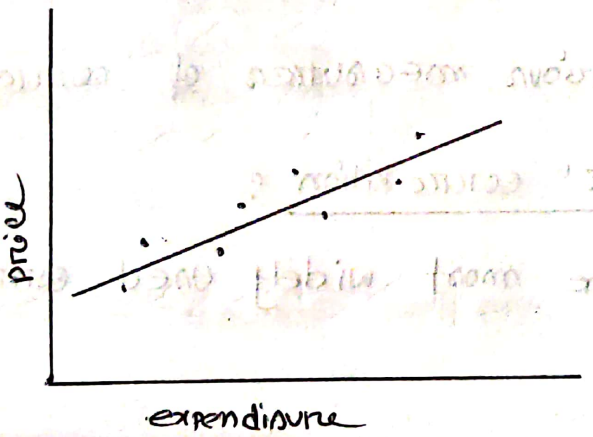


fig: positive correlation

**Negative correlation:** when two related variables move in opposite directions, their relationship is negative. When the coefficient of correlation "r" is less than 0 (zero), it is negative correlation. When r = -1, it is called perfect negative correlation. For example as the

stock prices rise, the bond market tend to decline.

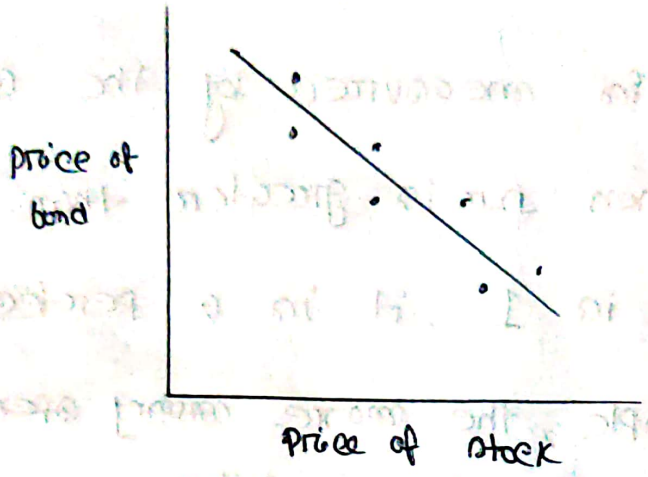


fig: Negative correlation.

There are various measures of correlation.

1) Pearson "r" correlation:

It is the most widely used correlation in statistic.

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}$$

2) Kendall rank correlation:

It is a non-parametic test that measures the strength of dependance between two variable.

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n-1)}$$

## Ans: to the que: NO - 3

(b)

Regression is a statistical method that attempts to determine the strength and the character of the relationship between one dependent variable (denoted by "y") and a series of other independent variable (denoted by "x")

### Objective of regression analysis:

It measures the average of dependent variable for the unit change of independent variable. Besides -

(i) Estimate the relationship between independent and dependent variable.

(ii) Determine the effect of each of the independent variables on the dependent variable.

(iii) Predict the value of the dependent variable for a given value of independent variable.