

(Short Questions)

① Statistics

The branch of science which deals with collection, presentation, analysis and interpretation of numerical data is called statistics.

Statistic

The numerical result which describes the characteristics of sample is called statistic.

② Give the functions/importance of statistics.

- Statistics assists in summarizing the larger sets of data in a form that is easily understandable.
- Statistics assists in the efficient design of laboratory and field experiments as well as surveys.
- Statistics assists in a sound and effective planning in any field inquiry.
- A modern administrator whether in public or private sector, leans on statistical data to provide a factual basis for decision.

③ Population

A collection of objects having some common characteristics of interest is called population.

Sample

A representative part taken from population is called sample.

④ What is inferential statistics?

The branch of statistics by which we draw a result about the characteristics of population using sample information taken from the population is called inferential statistics.

⑤ Primary data

The first hand collected data is called primary data.

- e.g., Email method, questionnaire method.

Secondary data

The data which has undergone in statistical process at least once is called secondary data.

- e.g., Punjab Bureau of Statistics, Pakistan Bureau of Statistics.

Write names of -
Measurement scales. By measurement we usually mean assigning of numbers to observations or objects and by scaling is a process of measurements or measuring.

- Nominal scale
- Ordinal Scale / Rank scale
- Interval Scale
- Ratio Scale

Classification

• The arrangements of objects into non-overlapping according to their similarities is called classification.

(7)

Tabulation

• The arrangements of objects into rows and columns according to similarities is called tabulation.

Frequency

• The total number of observations in to a group is called its frequency or group frequency.

(8)

Frequency distribution

• The allocation of total frequency into different classes or groups is called frequency distribution.

(9)

Give the types of average.

There are three types of average-

- Mean
- Median
- Mode

(10)

Give median with formula.

The value which divides the arranged data into two equal parts is called median.

• Formula:-

• For ungrouped data:

$$\text{Median} = \left(\frac{n+1}{2} \right)^{\text{th}}$$

• For grouped data:-

$$\text{Median} = l_b + \frac{h}{f} \left(\frac{n}{2} - c \right)$$

$\rightarrow l_b$ = lower class boundary of $\frac{n}{2}$ group

→ h = common interval

→ f = frequency of $\frac{n}{2}$ group

→ $\frac{n}{2}$ = indicating median group

→ C = cumulative frequency

(11)

What is empirical relationship?

When the distribution is moderately skewed then there exist relationship between the mode of distribution which is called empirical relationship.

(12)

What is dispersion?

The measure which calculates degree of spread the data from sum average value (mean, median, mode) ~~is~~ is called measure of dispersion. Dispersion means scatter / spread.

(13)

What is random experiment?

An experiment which produces different results, even it is repeated large numbers of time under identical condition is called random experiment.

(14)

What is event?

Any subset of sample space is called event e.g.-

$(S \sim)^A$ denote

$$S = \{1, 2, 3, 4, 5, 6\}$$

Let, $S \sim$ odd number

$$S = \{1, 3, 5\}$$

"S" is event

(15)

What is probability?

The chance of occurrence or not occurrence of an event is called probability. It can be further defined as the numerical measure of uncertainty is called probability.

(16)

(4)

What is frequency? Give its types.

• Frequency curve:-

When a frequency polygon or a histogram constructed over class intervals made sufficiently small for a large number of observations is smoothed, it approaches a continuous curve is called a frequency curve.

• Types:-

There are four types of frequency curve as follows:

- Symmetrical distribution.
- Moderately skewed.
- Extremely skewed.
- U-shaped distributions.

(17)

What is measure of dispersion?

In statistics, dispersion is the extent to which a distribution is stretched or squeezed - common examples of measures of statistical dispersion are the variance, standard deviation and interquartile range. There are two types of measure of dispersion: absolute and relative measure of dispersion.

(18)

What is standard deviation? give its properties.

The positive square root of the variance is called standard deviation. Symbolically:

$$S.D = \sigma = \sqrt{\frac{\sum (x_i - \bar{u})^2}{N}} \quad ; \text{ For population data}$$

$$S.D = \sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} \quad ; \text{ For sample data}$$

• Properties:-

- It is used to measure spread or dispersion around the mean of data set.
- It is never negative.

(19)

What is standard normal distribution?

The normal probability distribution of random variable which has zero mean and unit variance is called standard normal

distribution.

(20)

What is binomial experiment?

An experiment is called binomial experiment if it possess the following properties:

- The outcome of each trial may be classified into one of two categories, conventionally called success and failure.
- The probability of success, denoted by p , remain constant for all trials.
- The experiment is repeated a fixed number of times, say n .

(21)

What is expectation? give properties:-

Let X be discrete random variable which have possible values $x_1, x_2, x_3, \dots, x_n$ with corresponding properties $f(x_1), f(x_2), \dots, f(x_n)$, such that $\sum f(x) = 1$. Then the mathematical expectation or the expected value of X denoted by $E(X)$ is defined as:

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \dots + x_n f(x_n)$$

$$E(X) = \sum_{i=1}^{\infty} x_i f(x_i)$$

Properties:-

- The expected value of sum of any two random variables is equal to the sum of their expected values-i.e

$$E(x+y) = E(x) + E(y)$$

- The expected value of the product of two independent random variables is equal to the product of their expected values.i.e-

$$E(xy) = E(x) \cdot E(y)$$

(22)

What is random variable?

Such a numerical quantity whose value is determined by the outcome of a random experiment is called random variable.

(23)

Give properties of regression and co-relation.

• Properties of regression:-

- The least square regression line always goes through the point (\bar{x}, \bar{y}) the mean of data.

- The sum of the deviations of the observed values of y_i from the least square regression line is always zero. i.e. $\sum (y_i - \hat{y}_i) = 0$
- Properties of co-relation:-
- The co-relation coefficient lies between -1 and +1 i.e. $-1 \leq r \leq 1$
- The co-relation coefficient is independent of origin and scale.

(24)

Give the merits and demerits of arithmetic mean.

- Merits of arithmetic mean:-
- It is based on all the observations of data.
- It is determined for almost every kind of data.
- De-merits of arithmetic mean:-
- It is greatly affected by extreme values in the data.
- It sometimes gives fallacious conclusions.

(25)

Give the merits and de-merits of median.

- Merits of median:-
- It is easily calculated and understood.
- It is located even when the values are not capable of quantitative measurements.
- De-merits of median:-
- It is not rigidly defined.
- It is not capable of lending itself to further statistical treatment.

(26)

What is probability?

The word probability has two basic meanings: i) a quantitative measure of uncertainty and ii) a measure of degree of belief in a particular statement or problem.

(27)

Mutually exclusive events

- Two events A and B are said to be mutually exclusive if they both cannot occur at same time, so they have no points in common.

Non-mutually exclusive events

- Two events A and B are said to be non-mutually exclusive events if they both occur at some time, so they have some points in common.

(28)

Dependent events

The events A and B are said to be dependent if $P(A \cap B) \neq P(A) \times P(B)$. It means occurrence of one of events in some way affects the probability of the occurrence of the other event.

Independent events

The events A and B are said to be independent if the occurrence of any one event does not affect by the other event as:

$$P\left(\frac{A}{B}\right) = P(A) \text{ and } P\left(\frac{B}{A}\right) = P(B)$$

(29)

Give the law of addition for mutual events.

$$P(A \cup B) = P(A) + P(B)$$

(30)

Give the law of addition for non-mutual events.

$$P(A \cup B) = P(A) + P(B) + P(A \cap B)$$

It is often called as general rule of addition.

(31)

Simple hypothesis

A hypothesis is said to be a simple hypothesis if it specifies all the values of parameters of a distribution.

Composite hypothesis

A hypothesis is said to be a composite hypothesis if it does not specify the value of parameters of a distribution (inexact hypothesis).

(32)

Type I error \rightarrow False positive

If we reject a true null hypothesis then it is called type I error.

Type II error \rightarrow False negative

If we accept a false null hypothesis then it is called type II error.

(33)

Null hypothesis

The null hypothesis is a statement about the value of a population parameter.

Alternative hypothesis

A statement specifying that the population parameter is some value other than the one specified under the null hypothesis.

(34)

What estimation \Rightarrow estimator and estimate?

- Estimation:-

It is the process by which we get information about unknown values of population parameter by using sample values.

- Estimator: Estimator refers to that is used to generate estimate. It is a statistic that specifies how to use the sample data to estimate the unknown parameter of the population. It is a random variable.
- Estimate: Estimate refers to the fact or thing that we want to find out. The numerical value calculated from sample data is called estimate. It is number that is the computed value of estimation.

What is normal distribution?

(35) A normal distribution is a particular idealized, smooth, bell shaped histogram with all of the randomness removed. The normal distribution plays an important role in statistical theory and practice.

Give the properties of regression coefficient.

- If one regression coefficient is greater than one, then other will be less than one.
- The regression coefficient lies between $-\infty$ and $+\infty$.
- If x and y are independent random variables, then regression coefficient will be zero.

(36) Give the properties of co-relation coefficient.

- The correlation coefficient is symmetrical with respect to x and y and thus, $r_{xy} = r_{yx}$
- The correlation coefficient lies between -1 and $+1$.
- The correlation coefficient is zero when one of the variable x or y is constant.

(37) What is simple random sample?

A simple random sample is one in which every item from a population has the same chance of selection as any other item.

(38) Give the basic aim of sampling.

- To get minimum information about a population without

examining each and every unit of the population.

- To find reliability of estimates derived from the sample.

(40)

What is advantage of sampling?

- Sampling is cheaper than complete count.
- Sampling saves time.

(41)

Probability sampling

A probability sampling is one in which the sampling units are chosen on the basis of known probabilities.

Non-probability sampling

The sampling is said to be non-probability sampling when the procedure of selecting elements from the population is not based on probability but personal judgment is involved in selection.

(42)

What is level of significance?

The probability of rejecting a true null hypothesis is called level of significance.

(43)

What is sampling distribution of mean?

A probability distribution of all possible sample means of a given sample size is known as sampling distribution of mean.

(44)

Sampling error

The sampling error is the diff b/w a population parameter and a sample statistics.

Non-sampling error.

All types of errors other than sampling errors such as interviewer error and measurement error is called non-sampling error.

What is standard error?

(45)

The standard deviation of any estimator is called the standard error of the estimator.

(46)

What is random experiment?

An experiment which produces different results even though it is repeated a large number of times under essentially similar

(10) conditions is called a random experiment, e.g. the tossing of a fair coin, throwing of a balanced die.

Descriptive statistics

Descriptive statistics is that branch of statistics which deals with concepts and methods concerned with summarization and description of the important aspects of numerical data.

- e.g. A cricket player wants to find his score average for the last 20 games.

Inferential statistics

Inferential statistics deals with procedures for making inference about the characteristics that describes the large group of data or the whole called population, from the knowledge derived from only a part of data known as sample.

- A cricket player wants to estimate his chance of scoring based on his current season average.

Parameter

- Any numerical value which is calculated from population is called parameter.
- Parameters are not used to estimate population characteristics.
- e.g. $\mu \rightarrow$ mean (parameter)
 $\sigma \rightarrow$ S.D (parameter)
 $p \rightarrow$ Proportion

What is geometric mean?

(48)

Statistics

- Any numerical value which is calculated from sample is called statistics.
- Statistics are used to estimate population characteristics i.e. parameters.
- e.g. $\bar{x} \rightarrow$ mean (sample)
 $s^2 \rightarrow$ variance
 $s \rightarrow$ standard deviation
 $\hat{p} \rightarrow$ proportion

The geometric mean G of a set of n possible values x_1, x_2, \dots, x_n is defined as: the positive n th root of their product. i.e

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$$

where $x > 0$

when n is large, the computation of geometric mean becomes laborious as we have modify all the values and then extract the n th root.

(49)

Describe any two methods for collection of primary data.

- Direct personal investigation:

In this method an investigator collects the information personally.

(50)

from individual concerned.

- Indirect investigation:-

Sometimes the direct sources do not exist or information hesitate to respond for some reasons or others.

(51)

Give the disadvantages of range.

- It ignores all the information available from intermediate observation. As its value is based only on the two extreme observations, it might give a misleading picture of spread in the data.

(52)

Write any two properties of arithmetic mean.

- For a set of data, the sum of deviations of observations x_i 's from their mean \bar{x} , taken with these proper signs is equal to zero.

$$\begin{aligned} \text{The sum of deviation} &= \sum (x_i - \bar{x}) \quad (i = 1, 2, \dots, n) \\ &= \sum x_i - \sum \bar{x} \\ &= \sum x_i - n\bar{x} \\ &= 0 \end{aligned}$$

- If k subgroups of data consisting of n_1, n_2, \dots, n_k ($\sum n_i = n$) observations have respective mean $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$ then \bar{x} the mean of all data is given by:-

$$\begin{aligned} \bar{x} &= \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + \dots + n_k \bar{x}_k}{n_1 + n_2 + \dots + n_k} \\ &= \frac{\sum n_i \bar{x}_i}{n} \quad (i = 1, 2, 3) \end{aligned}$$

e.g. the weighted mean of all subgroups means.

- The sum of squared deviation of the x_i 's from mean \bar{x} is minimum.

(53)

Absolute dispersion

Absolute dispersion is one that measures the dispersion in terms of same units or in square of units as the units of data.

Relative dispersion

Relative dispersion is one that is expressed in the form of ratio, coefficient or percentage and is independent of units of measurement.

(54)

What is mutually exclusive event? give example..

(example #2)
When a die is rolled the outcomes are mutually exclusive as we get one ~~one~~ and only one of the six possible outcomes i.e. 1, 2, 3, 4, 5, 6 -

• (12) (die)

Two events A and B of a single experiment are said to be mutually exclusive or disjoint if and only if they can not both occur at some time. So, they have no points in common. e.g. when we toss a coin we get either a head or tail but not both. The two events Head and tail are therefore mutually exclusive. As: $A \cap B = \emptyset$ (null)

(55)

What is correlation?

Correlation is a measure of degree to which any two variables vary together.

(56)

What is ANOVA?

ANOVA stands for analysis of variance. ANOVA is a statistical method that separates observed variance data into different components to use the additional tests. One way ANOVA is used for 3 or more groups of data to gain information about the relationship between the dependent and independent variables.

(57)

What are class boundaries?

Class boundaries are the data values which separate classes. They are not part of class or data.

(58)

Define statistics in plural sense.

Statistics is a collection of methods that is used to analyze data.

Methods include collect, present and analyze data.

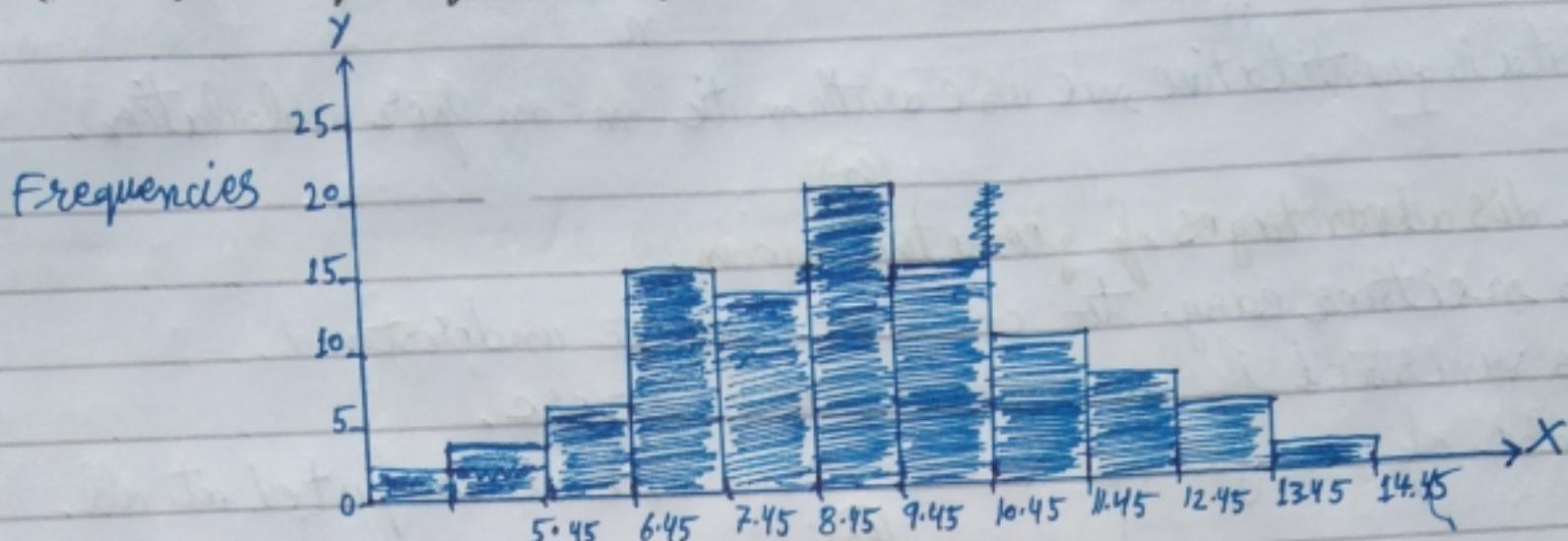
(59)

What is histogram?

A histogram consists of a set of adjacent rectangles whose bases are marked off by class boundaries (no class limits) and proportional to the frequencies associated with respective classes. The area of each rectangle represents the respective class frequencies. This is one of the most important graphical representation of a frequency distribution.

When the class intervals are equal, the rectangles have equal/ same width and their heights ~~are~~ directly represent the class frequencies; that is, they are numerically

proportional to the frequencies in the respective classes.
(Histogram for frequency distribution of annual death rates)



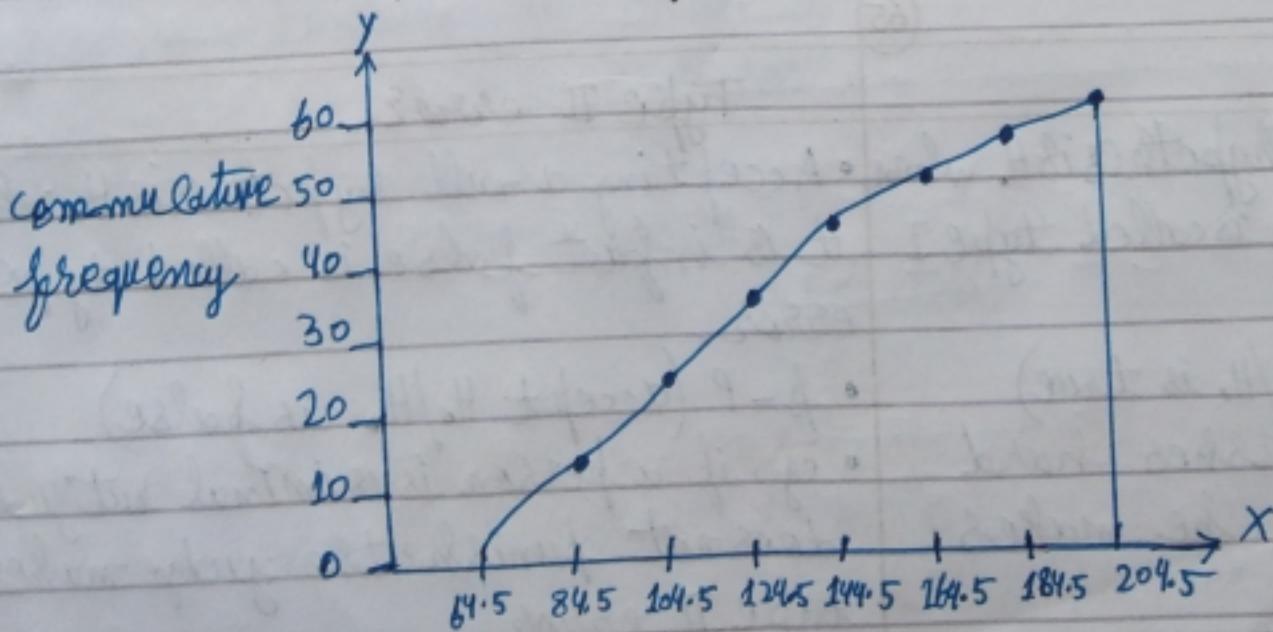
class boundaries -

⑥

Define cumulative frequency polygon or ogive?

A cumulative frequency polygon is popularly known as ogive (rhymes with alive and pronounced o-giv) is a graph obtained by plotting cumulated frequency of a distribution against upper and lower class boundaries depending upon whether the cumulation is of the "less than" or are joined by straight line segments because of its likeness to an architectural moulding called an ogel, a cumulative frequency polygon is called an ogive.

(Cumulative frequency Polygon (Ogive) for frequency distribution of weight of 60 apples



Upper class boundaries.

⑥

What is arithmetic mean? how it is calculated?

The sum of observations divided by the total number of observation

is called arithmetic mean. It is calculated as:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

When data is quantitative, we use arithmetic mean for calculation.

(62)

Write disadvantages of geometric mean.

- It is neither easy to calculate nor to understand.
- It is vanished if any observation is zero.
- In case of negative value, it can not be computed at all.

(63)

Write the properties of random experiment.

- The experiment can be repeated ~~only~~ any number of times.
- The experiment has always two or more possible outcomes.
- The outcomes of each repetition is unpredictable.

(64)

What is estimator? give example.

Estimator is a statistic that estimates some fact about the population. The rule that creates an estimate. e.g. - The sample mean (\bar{x}) is an estimator for the population mean (μ) i.e.

$$u \rightarrow \bar{x} = \frac{\sum x}{n} = \underbrace{70}_{\text{estimator}} \quad \underbrace{\text{estimate.}}_{\text{estimator}}$$

(65)

Type I error

- Rejecting a null hypothesis H_0 when it is in fact true is called type I error.

- $\alpha = P(\text{reject } H_0 / H_0 \text{ is true})$
- e.g. if teacher punishes hard working student, he makes type I error.

Type II error.

- Accepting a null hypothesis H_0 when it is in fact false is called type II error.

- $\beta = P(\text{accept } H_0 / H_0 \text{ is false})$
- e.g. if a person is criminal but judge does not punish, the judge makes type II error.

(66)

What is null hypothesis?

Any hypothesis which is to be tested for possible rejection under the assumption that is true. It should be always precise such as

(15)

If the given coin is unbiased or a drug is ineffective - on curing a particular disease. eg - suppose the average height of students in all colleges is 62" this statement is taken as hypothesis and is written symbolically as: $H_0: \mu = 62"$

(67)

What is alternative/research hypothesis? give example.

The hypothesis which we accept when the null hypothesis H_0 is rejected - It is ~~not~~ denoted by H_1 or H_A - H_0 is tested against H_1 eg. if our null hypothesis $H_0: \mu = 62"$, then our alternative hypothesis may be $H_1: \mu \neq 62"$ or $H_1: \mu > 62"$ or $H_1: \mu < 62"$

(68)

What is coefficient of correlation?

Pearson product moment correlation coefficient or coefficient of simple correlation is defined as: a numerical measure of strength of the linear relationship between any two variables is called Pearson product moment correlation coefficient or coefficient of simple correlation.

- Formula:-

$$\rho = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2(y - \bar{y})^2}}$$

$$-1 \leq \rho \leq +1$$

$\rho = +1$ perfect +ve correlation

$\rho = -1$ perfect -ve correlation

$\rho \rightarrow +1$ highly positive correlation

$\rho \rightarrow -1$ No correlation

$\rho = 0.30$ +vely correlated

$\rho = -0.33 \text{ to } -0.44$ -vely correlated

(69)

Give various relative measures of dispersion.

- Coefficient of dispersion/range:

$$\text{Range} = \frac{x_m - x_o}{x_m - x_o} \quad \because x_m \text{ is largest value \& } x_o \text{ is smallest value}$$

- Coefficient of quartile deviation:

$$Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} \quad \because Q_3 \text{ is upper quartile \& } Q_1 \text{ is lower quartile}$$

(Q16) (Reason)

• Coefficient of mean deviation:

$$= M.D \text{ (from mean)}$$

Mean

$$= M.D \text{ (from median)}$$

Median

∴ we use a relative measure of dispersion for comparing distribution of two or more data sets and for unit free comparison.

• Coefficient of standard deviation:

$$= \frac{S.D}{\text{mean}}$$

mean

• Coefficient of variance:

$$C.V = \frac{S.D}{\text{mean}} \times 100$$

70

Give the advantages of mean -

- It is rigorously defined by a mathematical formula.
- It is based on all the observations in data.
- It is amenable to mathematical treatment.

71

Finite population

A population is called finite if it is possible to count its individuals thus N is the size of the population

Infinite population

Sometimes it is not possible to count the units contained in the population such a population is called infinite or uncountable.

e.g. number of stars in universe, number of products in a market.

72

Find coefficient of range if $x_m = 100$ and $x_o = 25$.

$$\text{Coefficient of range} = \frac{x_m - x_o}{x_m + x_o}$$

$$= \frac{100 - 25}{100 + 25} = \frac{75}{125} = 0.6$$

73

What is ANOVA?

Analysis of variance (ANOVA) is a technique which partitions the total variation. A term distinct from variance and measured by the sum of squares of deviations from the mean

into its component parts, each of which is associated with a different sources of variations. These component parts of variance are then analysed in such a manner that certain hypothesis can be tested. This technique is based on the fact that the more the sample mean differ the larger the variance becomes.

(74)

Simple hypothesis

A simple hypothesis is one in which all parameters of the distribution are specified.

e.g.

If the heights of college students are normally distributed with $\sigma^2 = 4$ ($\sigma^2 = 4$) or $H: \mu = 62$ ($\sigma^2 < 4$) the hypothesis that its mean μ is say 62" that is $H: \mu = 62$. We have stated a simple hypothesis as the mean and variance together specify a normal distribution completely.

Composite hypothesis

A hypothesis which is not simple i.e all of the parameters of distribution are not specified is called composite hypothesis-

If we hypothesize that $H: \mu > 62$ ($\sigma^2 = 4$) or $H: \mu = 62$ ($\sigma^2 < 4$) the hypothesis becomes a composite hypothesis because we can not know the exact distribution of the population in either case. An indirect hypothesis is a composite hypothesis.

(75)

Parametric data

Parametric statistics are based on assumptions about the distribution of population from which the sample was taken.

Non-parametric data

Non-parametric statistics are not based on assumptions that is, the data can be collected from a sample that does not follow a specific distribution -

(76)

Primary data

Data that have been originally collected and have not undergone any sort of statistical treatment are called primary data.

Secondary data

Data that have undergone any sort of treatment by statistical methods at least once i.e data that have been collected, classified, tabulated or presented in some form for a certain purpose are called secondary data.

• It is past data

• It is real time data

- (18) ^{data}
- Primary sources are surveys, observations, experiments and personal interview etc.
 - Secondary data - sources include government publications books, articles and websites.

(77)

What is meant by classification?

The term classification is defined as the process of dividing a set of observations or objects into classes or groups in such a way that:

- Observations or objects in the same class or groups are similar.
- Observations or objects in each class or group are dissimilar to observations or objects in other class or group.

Classification is thus sorting of data into homogeneous class or groups according to their being alike or not.

(78)

Write disadvantages of mode:-

- It is vaguely defined.
- It is often intermediate and indefinite.
- It is not based on all the observations made.
- It is not capable of lending itself to further statistical treatment.

(79)

What are quantile points?

Quantiles are points in a distribution that relate to the ~~some~~ rank order values in that distribution. For a sample we can find any quantile by sorting the middle value of the sorted sample (middle quantile \rightarrow 50th percentile) is known as the median.

(80)

What is statistics?

Statistics is the branch of applied mathematics concerned with collecting, organizing and interpreting data. The data is represented by means of graphs. The two major areas of statistics are descriptive and inferential statistics.

(81)

What is level of significance?

The significance level of a test is the probability used as a

19

standard for rejecting a null hypothesis H_0 when it is assumed to be true. This probability is equal to some small pre-assigned value conventionally denoted by α . The value α is also known as the size of the critical region.

(82)

What is probability?

Probability is a mathematical tool used to study randomness. It deals with the chance (the likelihood) of an event occurring. e.g., if we toss a fair coin four times, the outcomes may not be two heads and two tails. However, if we toss the same coin 4000 times the outcomes will be close to half heads and half tails. The expected theoretical probability in any one toss is 1/2 or 0.5

(83)

Point estimate

When an estimate for the unknown population parameter is expressed by a single value, it is called as point estimate.

Interval estimate

An estimate expressed by a range of values within which the true value of the population parameter is believed to lie is referred to as interval estimate.

(84)

Dependent variables

- Dependent variable is a variable that can not be controlled or manipulated, its values are predicted from the independent variable.

- Dependent variables depend on other variables.

- They are explained by independent variables.

- They are the outcomes of research.

Independent variables

- Independent variable is a variable that can be controlled or manipulated.

- Independent variables do not depend on other variables.

- They explain the cause and effect of changes in the variable.

- They are exposure variable which are used in research.

(85)

What is regression model?

Regression model or regression analysis is a set of statistical methods

used to estimate relationship between a dependent variable and one or more independent variables. A linear regression line has an equation of the form:

$$Y = a + bX$$

Here X is the independent variable while " Y " is dependent variable.

(86)

Simple bar chart

- This chart consists of vertical or horizontal bars of equal width. The length of the bars represents the magnitude of the values of the variable i.e. the lengths of the bars vary depending on the size of data ~~not~~ values.

Multiple bar chart

- By multiple bar charts two or more sets of inter-related data are represented. The technique of simple bar chart is used to draw this chart but the difference we use different shades, colours or dots to distinguish between different phenomena.

(87)

What are non-mutually exclusive events?

Two events A and B are said to be mutually ^{non-}exclusive events if both the events A and B have at least one common outcome between them. e.g. In case of rolling a die the events of getting an odd face and the event of getting less than 4 are mutually non-exclusive events.

- Let " x " is denoted as event of getting an odd-face and " y " is denoted as event of getting less than 4. The event of getting less than 4. The event of getting an odd face = $x = \{1, 3, 5\}$

The event of getting less than 4 = $y = \{1, 2, 3\}$

- Between the events " x " and " y " the common outcomes are 1 and 3.

(88)

Give empirical relationship between mean, median and mode.

In statistics there is a relationship between ^{mean} median & mode which is empirically given by, the difference between the mean and the mode is three times the difference between the mean and the median. This relationship is given as:

$$\text{Mean} - \text{mode} = 3(\text{mean} - \text{median})$$