ALLIED COURSE - I

18K1SAS1

STATISTICS AND MATHEMATICS - I

UNIT - I

Census Method - Sampling method, Non - Probability sampling - Judgement Sampling, Quota sampling - advantages and disadvantages, Probability sampling - simple random sampling, Stratified random sampling, Systematic sampling - sampling errors.

UNIT - II

Vital statistics - Definition - Methods fertility - Crude birth rate, Specific birth rate, general birth rate, total fertility rate, gross reproduction rate and Net reproduction rate - Problems. Mortality - Crude death rate, SDR life table - uses - problems.

UNIT - III

Eigen Values and Eigen Vectors - Power of matrix, inverse of matrix - Cayley Hamilton theorem (without Proof) - Simple Problems.

UNIT - IV

Algebra - Binomial theorem, Expansion of rational fractions, summation of the series, approximation Exponential series - expansion - summation of the series, log with mic series - summation of the series - simple problems.

UNIT - V

Matrices - Definition, Types of matrices - operation on matrices, Hamilton matrix, orthogonal matrix. Rank of matrix. system of linear equations - Consistency - non-homogeneous linear equations, homogeneous linear equations - simple problems.

UNIT - I

SAMPLING DECISION

CENSUS METHOD

The object of a census or complete numeration is to collect information for each and every unit of the population is included in the investigation. When we make a complete enumeration of all items in the Population. It is known as census method of collection of data.

MERITS

* The data are collected from each and every item of the population.

* The results are more accurate and reliable because every item of the universe is enquired.

* Intensive study is possible.

* The data collected may be used for various surveys, analyses.

DEMERITS

* It requires a large number of enumerators and it is a really costly method. Therefore, the Government alone can use this method for conducting Population census, production census, etc.,

* It requires more money, labour, time, energy, etc,

* It is not possible in some circumstances where it the universe is infinite.

SAMPLING METHOD

According to Croxton and Cowdon,

It may be too-expensive or too-consuming to attempt either a complete or a nearly complete coverage in a statistical study. Further to arrive at valid conclusions, it may not be necessary to enumerate all or really nearly all of a population. We may study a sample drawn from the large population and if that sample is adequately representation of the population, we should be able to arrive at valid conclusions.

Snedecor observes, "A car load of coal is accepted or rejected on the evidences gained from testing only a few Pounds. The physician makes inferences about a patients blood through examination of a single drop.

Sample are devices for learning about large masses by observing a few individuals.

NON - PROBABILITY SAMPLING

a) JUDGEMENTAL SAMPLING

The investigation has the power to select or reject any item in an investigation. The choice of sample items depends on the judgement of the investigator. He has the vital role to play in collecting the information.

For example, if a sample of 5 students is to be selected from B.com class of 50 students for analysing the habit of picture seeing, the investigator would select 5 students who, according to his opinion are the representative of the class.

MERITS

* It is a simple method.

* It is used to uptain a more representative sample.

* It is very helpful to make public policies, decisions, etc., The executive and public officials use this method for their urgent problem.

DEMERITS

* Due to individual bias the sample may not to be representative one.

* It is difficult to get correct sampling errors.

* The estimates are not accuriate.

* It's results cannot be compared with others sampling studies.

b QUOTA SAMPLING

The sampling is similar to stratified Sampling. It is used in the U.S.A. for investigating Public opinion and consumer research.

To collect data, the universe is divided into quota according to some characteristics.

Each enumerator is then told to interview a certain number of persons who are his quota. The selection of sample items depends on personal Judgements.

It is a stratified cum - purposive sampling & thus has the advantages of both the methods. There is saving of time and money. If there are trained investigators, the sampling will give quite reliable results. Personal prejudice and individual bias are there. It is not based on random sampling, and so sampling error cannot be estimated.

PROBABILITY SAMPLING

According to Harper, " A random sample is a sample selected in such a way that every item in the population has an equal chance of being included.

(i) SIMPLE RANDOM SAMPLING

It is a technique in which sample is so drawn that each and every unit in the population has an equal and independent chance of being included in the simple.

a) Lottery method

b) Type of Random Sampling

a) LOTTERY METHOD

This is the most popular & simplest method. In this method, all the items of the universe are numbered on seperate slips of paper of same size, shape & colour. They are folded and mixed up in a drum or container. A blindfold selection is made. The required number of slips are selected for the desired sample size. The selection of item is depends on chance.

b) TYPE OF RANDOM NUMBERS :

As the lottery method cannot be used when the population is infinite the alternative method is that of using the table of random numbers.

There are several standard tables of random numbers . But the credit for this technique goes to prof. L.H.C. TIPPETT (1927) The random number table (taken from the British census Report) consists of 10,400 four figured numbers giving in all 10400 x 4 = 41600.

ii) STRATIFIED SAMPLING

When the population is heterogeneous or of different segments or strata with respect to the variable or characteristic under study, then it is stratified. First , the population is divided into a number of sub-groups or strata. Each stratum is homogeneous. A sample is drawn from each other stratum at random.

There are two types,

* Proportional
* Non-proportional

In proportional sampling equal and proportionate representative is given to sub-groups or strata. If the number of items is large in the population, the same will have a higher size and vice-versa.

In non-proportional sampling, equal representation is given to all the sub-strata regardless or their existence in the population.

(iii) SYSTEMATIC SAMPLING

It is also known as "QUASI-RANDOM SAMPLING". A systematic sample is selected at random sampling. When a complete list of the population is available, this method is used.

We arrange the items in numerical, alphabetical, geographical or any other order. K^{th} item is picked up from the sample frame.

$$K = N/n$$

K = sampling interval

N = size of universe

n = sample size

SAMPLING ERRORS

Sampling errors can be classified into two:

* Biased
* Unbiased

* BIASED ERRORS.

The errors that arise due to bias or pre-judice on the part of the informant or enumerator or investigator in selecting, estimating or measuring instruments are called biased errors. These errors are cumulative in nature and increase when the sample size also increases. These errors arise due to defect in the methods of collection of data, method of organisation of data and in method of analysis of data.

* UNBIASED ERRORS:

The errors, which arise in the normal course

of investigation or enumeration on account of chance, are called unbiased errors. They may arise accidently without any bias or pre-judice. These errors occur due to faulty planning of statistical investigation. To avoid these errors, the statistician must take proper precaution and case in using measuring instrument. He must see that the enumerators are also not biased. unbiased errors can be removed with the proper planning of statistical investigations. statisticians should have none of those errors.

MEASUREMENT OF ERRORS:

* ABSOLUTE ERROR = (ACTUAL VALUE) − (ESTIMATED VALUE)

$$\boxed{A.E. = a - e}$$

* RELATIVE ERROR = $\frac{a-e}{e}$

UNIT-II

VITAL STATISTICS

DEFINITION

Vital statistics include and mean the statistics of births and deaths.

Vital statistics deal mainly with the registration of birth & death and also provide many other information on the population growth. The births and deaths have an influence on the populations. The populations growth depends on births, deaths, marriage, divorce, seperation, rate of multiplicity, fertility rate, foundity of women, etc.,

MEASUREMENT OF FERTILITY

a) CRUDE BIRTH RATE :

Crude birth rate means number of persons, symbiocally,

$$\left.\begin{matrix}\text{Crude birth} \\ \text{rate}\end{matrix}\right\} = \frac{\text{Annual births} \times 100}{\text{Total population at mid-year}}$$

Crude birth rates are not generally suitable for the purpose of comparison. The reason is that the percentage of population in each age group may be different. In orderto make comparison of birth rates between two populations, allowances should be made for differences in ages and sex distribution. ie. the birth rate have to be standardised by a system of weighted arithmetic mean.

b) GENERAL FERTILITY RATE

The Total live births primarily depends upon the proportion of women in the child bearing age. Therefore the general fertility refers to the proportion of the number of children born per 1000 females (child bearing age). The formula is

$$GIFR = \frac{\text{NO. of live birth} \times 1000}{\text{NO. of female population of 15-49}}$$

PROBLEM:

Compute the general fertility rate from the data given below:

Age group (years)	No. of women (1000)	No. of live births
15 - 19	25	800
20 - 24	20	2400
25 - 29	18	2000
30 - 34	15	1500
35 - 39	12	500
40 - 44	6	120
45 - 49	4	10
TOTAL	100	7330

Out of 7330, the no. of female birth was 4000

$$\text{General fertility rate} = \frac{\text{No. of live birth}}{\text{No. of women in age [15-49]}} \times 1000$$

$$\text{G.F.R.} = \frac{7330}{1000000} \times 1000$$

G.F.R. = 7.33

c, SPECIFIC FERTILITY RATE :

The general fertility rate is an improvement over crude birth rate and it gives a general trend of the fertility rate of the child bearing age group, 15-49 years, as a whole. To have a detailed and better quality study, we have to calculate specific fertility rate for

different child bearing age groups. It is calculated as follows:

$$\left.\begin{matrix}\text{Age specific} \\ \text{fertility rate}\end{matrix}\right\} = \frac{\text{NO. of births of females of a specified age groups}}{\text{Mid year female population of a specific age-group}}$$

PROBLEM:

Compute the specific fertility rate,

Age group (years)	No. of womens (1000)	No. of live birth
15 - 19	25	800
20 - 24	20	2400
25 - 29	18	2000
30 - 34	15	1500
35 - 39	12	500
40 - 44	6	120
45 - 49	4	10
TOTAL	100	7330

out of 7330, number of female birth was 4000.

soln :-

$$SFR = \frac{\text{No. of live birth}}{\text{No. of womens}} \times 1000$$

$$SFR\ [15-19] = \frac{800}{25000} \times 1000 = 32$$

$$SFR\ [20-24] = \frac{2400}{20000} \times 1000 = 120$$

$$SFR\ [25-29] = \frac{2000}{18000} \times 1000 = 111.1\ldots$$

$$SFR\ [30-34] = \frac{1500}{15000} \times 1000 = 100$$

$$SFR\ [35-39] = \frac{500}{12000} \times 1000 = 41.67$$

$$SFR\ [40-44] = \frac{120}{6000} \times 1000 = 20$$

$$SFR\ [45-49] = \frac{10}{4000} \times 1000 = 2.5$$

D) TOTAL FERTILITY RATE :

When we find out the sum of age specific fertility rates at each age group say 15-49 years of age, it will give us the total fertility rate.

The formula is,

$$T.F.R. = \frac{SFR \times i}{1000}$$

PROBLEM :

Compute the TOTAL fertility rate from given below.

Age group (years)	No. of womens (1000)	No. of live births
15 - 19	25	800
20 - 24	20	2400
25 - 29	18	2000
30 - 34	15	1500
35 - 39	12	500
40 - 44	6	120
45 - 49	4	10
TOTAL	100	7330

Out of 7330, the number of female birth was 4000

Soln

Accordingly 1000 females exactly age 15, would by the time. They reach 20 have been $32 \times 5 = 160$ children.

It is necessary to multiply by 5.

EXACT Age	SFR x 5	Total birth Per 1000 female age 15 / started age
15	0	0
20	$35 \times 5 = 160$	160
25	$120 \times 5 = 600$	760
30	$111.1 \times 5 = 555.5$	1315.5
35	$100 \times 5 = 500$	1815.5

40	41.67 x5 = 208.35	2023.85
45	20x5 = 100	2123.85
50	2.5x5 = 12.5	2136.35

TOTAL FERTILITY RATE :

$$TFR = \frac{\Sigma(S.F.R. \times i)}{1000}$$

$$= \frac{2136.35}{1000}$$

$$= 2.136\,35$$

T.F.R. = 2.14

E) GROSS REPRODUCTION RATE :

The General fertility rate indicates the number of children born per 1000 women who are in 15-49 age group. The specific fertility rates indicates the number of children born to women of a particular age group, say 15-49, Per thousand, these fertility rates are unsuitable for the study of the growth of population. The population increases mainly through female babies, who are future mothers. In order to have an idea of the reproduction of population, we have to take into account only the number of babies which is an important factor.

General reproduction rate indicates the

average number of female children born to a women of child bearing age.

For e.g.

100 females give birth to 100 females G.R.R. = 1

100 females give birth to 150 females G.R.R. = 1.5

100 females give birth to 80 females G.R.R. = 0.8

If G.R.R. is one, it indicates that the set under consideration is exactly replacing. If the rate is more than one, the population would increase.

$$G.R.R. = \frac{\text{No. of female birth} \times \text{Total F.R.}}{\text{Total number of birth}}$$

The G.R.R. does not take into account the mortality factor, but shows rate at which mothers would be replaced by daughters without considering death rate.

PROBLEM:

Calculate the Gross Reproduction Rate from the data.

Age group	Female Population (1000)	Female boxins	survival rate
15-19	1600	19000	0.921
20-24	1000	70200	0.901
25-29	1685	90600	0.885
30-34	1730	62400	0.862
35-39	1725	32500	0.850
40-44	1620	11000	0.832
45-49	1510	800	0.812

Soln :-

$$\text{S.F.R. Per women} = \frac{\text{No. of female birth}}{\text{Female population}}$$

Age	Female population	Female born	S.F.R. Per women
15-19	1600000	19000	0.011875
20-24	1000000	70200	0.0702
25-29	1685000	90600	0.05438
30-34	1730000	62400	0.0361
35-39	1725000	32500	0.019
40-44	1620000	11000	0.0068
45-49	1510000	800	0.00053
			= 0.1981

$$N.R.R. = \frac{\Sigma \left(\begin{matrix} \text{No. of female} \\ \text{children born} \end{matrix} \right) \times \left(\begin{matrix} \text{No. of} \\ \text{survival rate} \end{matrix} \right)}{1000}$$

$$= \frac{1149}{1000}$$

$$N.R.R. = 1.149$$

MEASUREMENT OF MORTALITY:

a) GENERAL DEATH RATE (CRUDE DEATH RATE)

The general, death rate is widely used, because it is easy to calculate and to understand. It requires only the total population and the total number of deaths. It gives preliminary indication of the level of mortality.

The crude death rate measures the decreases in population due to death. crude death rate is not suitable for comparing the mortality Prevailing in different regions, because the composition of different age groups may differ widely. Therefore, Standard death rates are to be calculated for purpose of comparison.

PROBLEM

Compute the Crude death rate.

Gross reproduction rate = $\leq SFR \times 5$

$= 0.1981 \times 5$

G.F.R. $= 0.9905$

F) NET REPRODUCTION RATE:

G.R.R. indicates the number of female children born to 1000 women of child bearing age. It means and indicates the average number of female children born per woman of child-bearing age. G.R.R. does not take into account the current mortality rate. N.R.R. is nothing but G.R.R. adjusted for the effects of mortality rate. N.R.R. is nothing It measures the extend to which generation of girl babies survive to reproduce babies of their sex, as they pass through the child bearing age group. The formula is,

$$N.R.R. = \frac{\text{No. of female children born \& survivors to 1000 females}}{1000}$$

N.R.R. is a good measure of population growth. It can never be greater than G.R.R. because N.R.R. taken into account the mortality factor.

PROBLEM:

Calculate the Net Reproduction from the given data:

Age group of child bearing female	No. of female children born 1000 women passing through age/ age group	No. of survival rate out of each 1000 female children.
15 - 20	50	~~380~~ 850
20 - 25	180	800
25 - 30	450	750
30 - 35	500	700
35 - 40	300	650
40 - 45	100	600
45 - 50	40	500

Soln

$$\left.\begin{matrix}\text{No. of survival rate}\\ \text{which replaced present}\\ \text{women}\end{matrix}\right\} = \frac{\left(\begin{matrix}\text{No. of female}\\ \text{child born}\end{matrix}\right) \times \left(\begin{matrix}\text{No. of survival}\\ \text{female}\end{matrix}\right)}{1000}$$

Age group	No. of female children born 1000 women Passing through	No. of survival	No. of survival rate which replaced
15 - 20	50	850	42.5
20 - 25	180	800	144
25 - 30	450	750	387.5
30 - 35	500	700	350
35 - 40	300	650	195
40 - 45	100	600	60
45 - 50	40	500	20

Standardised death rate (TOWN A)

$$= \frac{[(15000 \times 24) + (20000 \times 20) + (10000 \times 28)]}{15000 + 20000 + 10000}$$

$$= \frac{[360000 + 400000 + 280000]}{45000}$$

$$= \frac{1040000}{45000}$$

(A) = 23.11

Standardised death rate (TOWN B)

$$= \frac{[(40000 \times 25) + (52000 \times 20) + (8000 \times 30)]}{40000 + 52000 + 8000}$$

$$= \frac{[1000000 + 1040000 + 240000]}{100000}$$

$$= \frac{2280000}{100000}$$

(B) = 22.8

We can say that the death rate is town B is higher than in town A.

UNIT - III

EIGEN VECTORS AND EIGEN VALUES

Suppose we have a vectors $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ where $x_1^2 + x_2^2 + x_3^2 \neq 0$ if we are transforming a vectors x by means of matrix A.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$ We get a vector

$$Y = AX \longrightarrow (1)$$

Parallel to the vector x, we have $y = \lambda x \rightarrow (2)$

[λ is a scalar] from (1) and (2) wehave $Ax = \lambda x \rightarrow (3)$

Then the vector x is tesmet the Eigen vector of matrix A or the Eigen vectors of the given linear transformation the scalar λ is tesmet the Eigen value associated with this Eigen vectors. Other names for eigen vectors are characteristic vectors, Proper ~~wa~~ Vector and atent vectors. Eigen Values are also know as characteristic vectors proper vectors and Roots.

PROBLEM:

Find the eigen Values and eigen Vectors at the matrix $A = \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix}$.

The characteristic equation is

$$|A - \lambda I| = 0$$

$$|A - \lambda I| = 0$$

$$\begin{vmatrix} 4-\lambda & 1 \\ 3 & 2-\lambda \end{vmatrix} = 0$$

$$(4-\lambda)\ (2-\lambda) - 3 = 0$$

$$8 - 4\lambda - 2\lambda + \lambda^2 - 3 = 0$$

$$8 - 6\lambda + \lambda^2 - 3 = 0$$

$$-6\lambda + \lambda^2 + 5 = 0$$

$$\lambda^2 - 6\lambda + 5 = 0$$

(using Quadratic equation)

$$(\lambda - 1)\ (\lambda - 5) = 0$$

$$\boxed{\lambda = 1} \quad \boxed{\lambda = 5}$$

The eigen values are 1 and 5 two fine the eigen vectors,

when $\lambda = 1$ $|A - \lambda I|\, x = 0$

$$\begin{vmatrix} 4-\lambda & 1 \\ 3 & 2-\lambda \end{vmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\begin{vmatrix} 4-1 & 1 \\ 3 & 2-1 \end{vmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\begin{vmatrix} 3 & 1 \\ 3 & 1 \end{vmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$3x_1 + x_2 = 0$$

$$\boxed{x_1 = K}$$

$$x_2 = -3x$$

$$\boxed{x_2 = -3K}$$

Note that whatever we the value of x_1, the value of x_2 is 3 times of that, $\therefore$ the eigen vectors corresponding to $\lambda = 1$ is $\begin{bmatrix} K \\ -3K \end{bmatrix}$ which is the general form where K is a constant.

Hence, we get infinite number of eigen vectors by giving different value for K.

The simplest eigen vector is,

$$\therefore \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

$\lambda = 5$ $\quad |A - \lambda I| x = 0$

$$\begin{bmatrix} 4-\lambda & 1 \\ 3 & 2-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} -1 & 1 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$-x_1 + x_2 = 0 \rightarrow +x_1 = +x_2$$

$$3x_1 - 3x_2 = 0 \rightarrow x_1 = \frac{3x_2}{3}$$

$$\boxed{x_1 = K}$$

$$x_1 = x_2 \quad , \quad \boxed{K = x_2}$$

By giving $x_2 = K$, we have $x_1 = K$ and therefore for different values of K, we get different eigen vectors. The general eigen vector is :.

Hence $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ is an eigen vectors corresponding to $\lambda = 1$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is eigen vectors corresponding to $\lambda = 5$.

PROPERTIES OF EIGEN VECTORS

THEOREM 1 :

Statement : Let $\lambda_1, \lambda_2 \ldots \lambda_n$ by the eigen values corresponding to a square matrix. A or order n and $x_1, x_2, \ldots x_n$ or corresponding vectors.

Then $x_1, x_2, \ldots x_n$ or linearly independent. That is eigen vectors corresponding to eigen value are linearly independent.

THEOREM 2 :

If x is a characteristic vector of a matrix A corresponding to the characteristic root λ then Kx for every non-zero scalar K is also a characteristic vector of A corresponding to λ.

Note 1.

This theorem implies that, to one characteristic root we have many eigen vectors.

Note 2.

The above theorem may also be stated as follows,

"corresponding to a characteristic root λ of a there are different eigen vectors."

THEOREM 3:

A eigen vectors x of a matrix A cannot correspond to more than one eigen root of A.

PROPERTIES OF EIGEN VALUES

1) Some of eigen values is equal to the some of the diagonal elements.

2) products of eigen values is equal to it determinant value.

Note,

If $\lambda_1, \lambda_2, \ldots \lambda_n$ be eigen values A, then $1/\lambda_1, 1/\lambda_2, \ldots 1/\lambda_n$ will be eigen values of A^{-1}.