

UNIT 3 Statistics part

Syllabus : Elements of Statistical methods Variables, Distributions: Primary data and secondary data, Population and sample, Census and Sample Surveys. Tabulation of data and Diagrams, Graphs, Bar Diagram, Pie Diagram etc. Frequency ~~Tables~~ Distributions, Ungrouped and grouped, continuous frequency distribution. Histogram, Inequality curve, Measures of Central tendencies, Averages: \sum , GM, HM, Mean, Median and Mode (Their advantages and disadvantages) Measures of Dispersion - Range, Quartile Deviation, Mean Deviation, Variance / S.D., Moments, Skewness and Kurtosis.

Sampling theory : Meaning and objects of sampling. Some ideas about the methods of selecting samples, ~~Statistic~~ ^{Statistic} and parameter, Sampling Proportion. Four fundamental distributions derived from the normal: (i) Standard Normal Distribution, (ii) Chi-square distribution (iii) Student's t -distribution, (iv) Smeedecor's F-distribution. Estimation and Test of significance. Statistical Inference. Theory of Estimation. Point estimation and Interval estimation. Confidence Interval / Confidence Limit. Statistical Hypothesis - Null Hypothesis and Alternative Hypothesis. Level of Significance. Critical Region. Type I and II error Problems. Bivariate Frequency Distribution. Scatter Diagram. Correlation coefficient. Definition

and properties of regression lines.

Books followed: Statistical Tools & Techniques - P. K. Ghiri & J. Banerjee

1. Fundamentals of Statistics - A. W. Goodman, M. K. Gupta, B. Dasgupta

2. Statistical Methods - N. G. Das

1. Statistics: In modern age, the word statistics includes all branches of human activity concerned with numerical data.

Statistics is used as a plural noun and also a singular noun.

In plural sense, numerical data, i.e., numerical statements of facts in connection with any field of human experience are called statistics. For instance, production statistics of a factory over different years, statistics of birth and death of a country, etc.

The term statistics, in singular sense, means the aggregate of scientific methods which deals with the collection, presentation, analysis and interpretation of numerical data.

Statistical Methods are found to be applicable in several fields, namely, economics, business and industry, education, psychology, biology, agriculture etc.

1.1. Statistical data: Primary data and Secondary data.

Statistical data are classified as primary and secondary, depending on the source of the data.

(i) Primary data: The statistical data which are gathered

Math (GATE)

GE4 (SB)

directly from the field of investigation for the desired purpose are called primary data. It is obvious that these data are original in nature.

A doctor, interested in the weights of his patients, records their weights using a machine. These data are primary to the doctor.

This type of data can be used with greater confidence, because the enquirer himself decides upon coverage of the data, the definitions used, the method of collection of the data, etc. and as such, he has a measure of control on the reliability of the data.

But collection of such data requires more money, man-power and time.

(ii) Secondary data: The statistical data which have already been collected by some agency are compiled from that source by the enquirer for his use are called secondary data. In other words, data collected by one when used by another, or collected for one purpose when used for a different one, are termed as secondary data. It is evident that secondary data are not original.

The census data collected from census reports by a research scholar for his study are secondary data to the scholar.

Collection of secondary data is cheaper as it involves less

man-power and time.

But secondary data usually contain error due to ~~transcription~~ transcription, rounding, etc. and hence are hardly reliable. In making use of such data, the enquirer has to be particularly careful about the coverage of the data, the definitions on which they are based, method of collection of data and their degree of reliability.

From the above ~~discussion~~ discussion, it is clear that the same data which are primary for one may be secondary for some other persons.

1.2 Attribute and Variable

Statistical data are numerical, although in some situations, data may not be quantitative at first stage. At the outset these non-numerical data are appropriately classified under different mutually disjoint categories and, subsequently, number of items included in each category is obtained.

An attribute is a qualitative character that can not be numerically expressed. Individuals possessing an attribute can, however, be grouped into several disjoint classes. Mother tongue of a group of people, grade obtained by students in a test or colour of flowers represents an attribute.