

Department of Statistics
Discipline Specific Minor Course
(Semester-II)

Year	Semester	Course Type	Course Code	Course Title	Theory/ Practical	Credits	No. of Theory/ Practical to be conducted	Page No.
1 st	II	DSC (Minor)	ST-MN-126T	An Introduction to Statistical Learning	Theory	2	30L	2-4

DISCIPLINE SPECIFIC MINOR COURSE (ST-MN-126T):
An Introduction to Statistical Learning

Course Code & Title	Credits	Credit Distribution of the Course	
		Theory	Practical
ST-MN-126T - An Introduction to Statistical Learning	2	2	---

LEARNING OBJECTIVES:

The Learning Objectives of this course are as follows:

- To acquaint students with some basic concepts in Statistics.
- To understand the concept of primary data, secondary data, data collection methods, statistical population and sampling methods.
- To compute various measures of central tendency and measures of dispersion.
- To compute raw and central moments and various measures of skewness and kurtosis.

COURSE OUTCOMES:

After completion of this course student will able to;

CO-1: Understand meaning and importance of statistics.

CO-2: Understand the concept of primary data, secondary data, data collection methods, statistical population and sampling methods.

CO-3: Compute various measures of central tendency and measures of dispersion.

CO-4: Compute raw and central moments and various measures of skewness and kurtosis.

SYLLABUS OF ST-MN- 126T**[30 Hours]****UNIT-I: Introduction to Statistics****[2 Hours]**

1.1 Meaning of Statistics as a Science.

1.2 Importance of Statistics.

UNIT-II: Population and Sample**[8 Hours]**

2.1 Types of characteristics: Attributes: Nominal scale, ordinal scale, Variables: Interval scale, ratio scale, discrete and continuous variables,

2.2 Types of data: primary data, secondary data, cross-sectional data, time series data, directional data.

2.3 Data collection methods (Survey, laboratory experiments, simulation), concept of big data, role of statistics in data science. Concept and types of statistical population: Finite

population, infinite population, homogeneous population and heterogeneous population.

- 2.4** Concept of a sample and Methods of sampling: Probability sampling, Simple random sampling with and without replacement (SRSWR and SRSWOR), stratified random sampling, systematic sampling, cluster sampling and two-stage sampling (only Description).
- 2.5** Non probability sampling: Judgment sampling, Quota sampling, Convenience and Snowball sampling.
- 2.6** Frequency Classification: Raw data and its classification, ungrouped frequency distribution, Sturge's rule, grouped frequency distribution, cumulative frequency distribution, inclusive and exclusive methods of classification, Open end classes, and relative frequency distribution.

UNIT-III: Summary Statistics:

[13 Hours]

3.1 Measures of Central Tendency:

Concept of central tendency of statistical data, Statistical averages, characteristics of a good statistical average. Arithmetic Mean (A.M.): Definition, Properties of A.M, combined mean of a number of groups, merits and demerits, trimmed arithmetic mean. Mode and Median: Definition, formulae (for ungrouped and grouped data), merits and demerits. Empirical relation between mean, median and mode. Partition Values: Quartiles, Deciles and Percentiles (for ungrouped and grouped data), Box Plot. Geometric Mean (G.M.): Definition, merits and demerits. Harmonic Mean (H.M.): Definition, merits and demerits. Order relation between arithmetic mean, geometric mean and harmonic mean. Weighted A.M.

3.2 Measures of Dispersion:

Concept of dispersion, characteristics of good measure of dispersion, Range, coefficient of range, Quartile deviation (Semi-interquartile range): Definition, merits and demerits, coefficient of quartile deviation. Mean deviation: Definition, merits and demerits, minimal property (without proof), coefficient of mean deviation. Mean squared deviation: Definition, minimality property of mean squared deviation (with proof). Variance and standard deviation: Definition, merits and demerits, Properties of variance, combined variance for n groups (derivation for two groups). Coefficient of variation (C.V.)

UNIT IV: Moments, Skewness and Kurtosis:

[7 Hours]

- 4.1** Raw moments for ungrouped and grouped data. Central moments for ungrouped and grouped data, Effect of change of origin and scale. Relations between central moments and raw moments, up to 4-th order (without proof).
- 4.2** Concept of skewness of frequency distribution, positive skewness, negative skewness, symmetric frequency distribution. Bowley's coefficient of skewness: Bowley's coefficient of skewness lies between -1 to 1 (with proof), interpretation using Box plot. Karl Pearson's coefficient of skewness. Measures of skewness based on moments.
- 4.3** Concepts of kurtosis, leptokurtic, mesokurtic and platykurtic frequency distributions.

Measures of kurtosis based on moments.

ESSENTIAL / RECOMMENDED READINGS:

- 1) Agarwal, B. L. (2003). Programmed Statistics, Second Edition, New Age International Publishers, NewDelhi.
- 2) Ghosh, J. K. and Mitra, S. K., Parthsarthy, K. R. (1993). Glimpses of India's StatisticsHeritage, Wiley publishing Co.
- 3) Goon, A. M., Gupta, M. K. and Dasgupta, B. (1983). Fundamentals of Statistics, Vol. 1, Sixth Revised Edition, The World Press Pvt. Ltd., Calcutta.
- 4) Gupta, S. C. and Kapoor, V. K. (1983). Fundamentals of Mathematical Statistics, Eighth Edition, Sultan Chand and Sons Publishers, NewDelhi.
- 5) Gupta, S. C. and Kapoor, V. K. (1997). Fundamentals of Applied Statistics, Third Edition, Sultan Chand and Sons Publishers, NewDelhi.
- 6) Neil A. Weiss, (2016). Introductory Statistics, Tenth Edition, Pearson.
- 7) Purohit, S. G., Gore S. D., Deshmukh S. R. (2008). Statistics Using R, Narosa Publishing House, NewDelhi.
- 8) Sarma, K. V. S. (2001). Statistics Made it Simple: Do it yourself on PC. Prentce Hall of India, NewDelhi.
- 9) Snedecor G. W. and Cochran W. G. (1989). Statistical Methods, Eighth Ed. East-West Press.

