Statistics

4 year B.Sc.

Year	Subjects	Theory/Practical	Full Marks	Total
I	Fundamentals of Statistics (STA 101)	Theory	100	
	Fundamentals of Statistics (STA 102)	Practical	50	150

Tribhuvan University
Institute of Science & Technology

Course Title: Fundamentals of Statistics — Full Marks: 100
Course Code: STA 101 — Pass Marks: 35
Level: B.Sc. — Total Number of Periods: 150
Year: I
Nature of the Course: Theory

Course objectives:

To impart the knowledge of descriptive statistics, correlation, regression, theoretical as well as the applied knowledge of probability and some probability distributions

Group A

1. Introduction to Statistics **[5]**

Meaning of Statistics as a Science; Importance of Statistics; Scope of Statistics in the field of physical Sciences, Biological Sciences, Medical Sciences, Industry, Economics Sciences, Social Sciences, Management Sciences, Information Technology, Agriculture, Insurance, Education and Psychology.

2. Population and Sample **[8]**

Types of Characteristics; Scales of measurement; qualitative, quantitative, discrete and continuous variables, entities; Types of Data: (i) primary data, secondary data and their sources (ii) cross-sectional data, time series data, failure time data, panel data; Notion of a statistical population: finite population, infinite population, homogeneous population and heterogeneous population; Notion of sample, random sample and non-random sample; methods of sampling (description only): simple random sampling with replacement (SRSWR) and without replacement (SRSWOR).

3. Presentation of Data [6]

Organization of Data: Data mining, editing, coding and data management; assessing the quality of the data; Classification and Tabulation : Raw data and its classification, Discrete frequency distribution, construction of class interval (Sturge's rule), continuous frequency distribution, inclusive and exclusive methods of classification, open end classes, cumulative frequency distribution and relative frequency distribution; tabulation, construction of bivariate frequency distribution. Diagrammatic Presentation of Data: Simple bar diagram, multiple bar diagram, sub-divided bar diagram, pie-chart (review). Graphical Presentation of Data: Histogram, frequency curve, frequency polygon, ogive curves stem and leaf chart, range chart; Check sheet, Pareto diagram

Problems and illustrative examples

4. Measures of Central Tendency and Dispersion [17]

Concept of measures of central tendency; mathematical properties of arithmetic mean, weighted arithmetic mean, trimmed mean, formula for computation of mode and median (with derivation) graphical method, harmonic mean, weighted harmonic mean geometric mean, weighted geometric mean, order relation between arithmetic mean, geometric mean, harmonic mean (proof for n = 2), problems focusing on theoretical aspects, empirical relationship between mean , median and mode, choice of appropriate average

Concept of measures of dispersion, different methods of measuring dispersion, absolute and relative measures of dispersion, minimality property of mean deviation, minimality property of mean square deviation(with proof), variance and standard deviation, mathematical properties of standard deviation, effect of change of origin and scale in standard deviation, combined variance(derivation for 2 independent groups), generalizations for n groups, coefficient of variation(C.V.), theoretical problems of measures of dispersion; empirical relationships, five number summary; box plot, normal probability plot; Lorenz curve, Ginni coefficient

Problems and illustrative examples

5. Moments, Skewness and Kurtosis [12]

Raw moments (m_r') for grouped and ungrouped data; moments about an arbitrary constant for grouped and ungrouped data $m_r(a)$; Central moments (m_r) for grouped and ungrouped data; Effect of change of origin and scale; Relations between central moments and raw moments (up to 4^{th} order).

Concept of skewness of frequency distribution; positive skewness, negative skewness, symmetric frequency distribution, Bowley's coefficient of skewness : Computation of coefficient of skewness using Bowley's formula and its interpretation, interpretation using Box plot; Karl Pearson's coefficient of skewness; Measures of skewness based on moments (β_1, γ_1), Concepts of kurtosis; leptokurtic, mesokurtic and platykurtic frequency distributions; measures of kurtosis using partition values; Measures of kurtosis based on moments (β_2, γ_2)

Problems and illustrative examples

6. Introduction to Correlation [12]

Bivariate data, bivariate frequency distribution, correlation between two variables, positive correlation, negative correlation, scatter diagram to explore the type of correlation, covariance between two variables: Definition, computation, effect of change of origin and scale; Karl Pearson's coefficient of correlation (r): Definition, computation for grouped and ungrouped data and interpretation, assumptions for Karl Pearson's correlation coefficient, theoretical problems

Properties (with proof): (i) $-1 \leq r \leq 1$, (ii) Effect of change of origin and scale

Spearman's rank correlation including tied cases

Problems and illustrative examples

7. Regression Analysis [15]

Concept of regression, lines of regression, fitting of lines of regression by the least squares method, interpretation of slope and intercept, concept of linearity

Regression coefficient (b_{yx}, b_{xy}): Definition, computation, properties (with proof).

(i) $b_{yx}b_{xy} = r^2$, (ii) $b_{yx}b_{xy} \leq 1$, (iii) $b_{yx} = r\frac{\sigma_y}{\sigma_x}$, $b_{xy} = r\frac{\sigma_x}{\sigma_y}$,

(iv) Effect of change of origin and scale, (v) Angle between the two lines of regression

Mean residual sum of squares, Residual plot and its interpretation for assessing the goodness of fit of the regression line, explained and unexplained variation, coefficient of determination; concept of multiple regression

Problems and illustrative Examples

Group B

8. Introduction to Probability [20]

Review of set operations; Concepts in probability: deterministic and random experiments; Definitions of terms: trial and event, outcome, sample space, equally likely, mutually exclusive, exhaustive and favorable cases, sure and impossible events, independent and dependent events; Definitions of probability: mathematical (classical), statistical (relative frequency) and subjective with their merits and demerits; Combinatorial analysis and combinatorial probability examples, algebra of events and probability; Properties of probability and basic theorems: Additive and multiplicative theorems, Boole's inequality; Axiomatic definition of probability, geometrical probability and Bertrand's paradox; Conditional probability, pair-wise and mutual independence, Bayes theorem, prior and posterior probabilities, sensitivity, specificity, predictive value positive and predictive value negative of a diagnostic test

Problems and illustrative examples

9. Random Variables **[10]**

Concept of a random variable, types of random variables: Discrete and continuous random variables; Probability distribution of a random variable: probability mass function and probability density function, distribution function and its properties; Functions of random variables, examples of linear and nonlinear transformations.

Problems and illustrative examples

10. Theory of Mathematical Expectation **[15]**

Mathematical expectation of a random variable (discrete and continuous) and its function, properties of mathematical expectation of random variables, addition and multiplicative theorems of expectation, covariance and correlation, conditional expectation, conditional variance, variance of a linear combination of random variables; Moments of random variables: Raw and central moments, uses of moments, obtaining measures of location (averages), dispersion, skewness and kurtosis of a given probability distribution; Generating functions: Moment generating function, probability generating function, cumulant generating function and characteristic function with their properties.

Problems and illustrative examples

11. Probability Distributions **[30]**

Discrete distributions: Bernoulli trial, binomial and Poisson distributions, their mass functions, distribution functions, moment generating functions, characteristic functions, moments, properties, distribution fittings; Continuous distributions: Rectangular and normal distributions: their probability density functions, distribution functions, moment generating and characteristic functions, properties and uses, normal distribution as an approximation of binomial and Poisson distributions, standard normal distribution, distribution fittings.

Problems and illustrative examples

References:

1. Miller and Fruend (2007). *Modern Elementary Statistics*, Pearson Publishers.
2. Snedecor and Cochran (1980). *Statistical Methods*, Oxford and IBH Publishers
3. Gupta S.C. and Kapoor V.K.(2012). *Fundamentals of Mathematical Statistics,* Sultan Chand and Sons, New Delhi
4. Shrestha H.B., *Statistics and Probability: Concepts and Techniques*, EKTA Books. Latest Edition.
5. Sthapit Azaya, Yadav Rashinder, Khanal Shankar. (2012). *Fundamentals of Statistics*, Asmita Publication, Kathmandu, Nepal
6. *Sukubhattu*, N. P. (2063 BS). *Probability Theory and Statistical Methods*, 2nd edition, Asmita Publications, Kathmandu
7. Daniel, W.W (2000). *Biostatistics: A foundation for analysis in the health sciences*, 7th edition, John Wiley and sons, INC

Tribhuvan University
Institute of Science & Technology

Level: B.Sc.
Year: I
Course Title: Fundamentals of Statistics
Course Code: STA 102 Full Marks: 50
Nature of the Course: Practical Pass Marks: 20
Total Number of Periods:180

Pre-requisites: Knowledge of the topics in theory, and the laboratory with well-equipped computers facility should be arranged..

Course objectives:

- To develop computational skills in descriptive statistics and probability
- To apply theoretical knowledge in practical numerical problems
- To make students familiar for handling statistical software for data analysis

Practical problems

S.No.	Title of the practical problems	No. of practical problems
1	Arrangement of raw data pertaining to discrete and continuous variables into the proper format for further statistical analysis using appropriate codes (if necessary) (Also use MS EXCEL Spread sheet)	1
2	Preparation of frequency distribution, cumulative frequency distribution, histogram, frequency curves, stem and leaf plot, box and whisker plot(Also use MS EXCEL Spread sheet and any statistical software such as SPSS, STATA etc. whichever convenient)	1
3	Diagrammatical presentation of data(Also use MS EXCEL Spread sheet) with problems based on simple diagram, subdivided bar diagram, Pie diagram etc.	1
4	Problems using Pareto Diagram	1
5	Computation of measures of central tendency (ungrouped and grouped data) Use of an appropriate measure and interpretation of results and computation of partition values (Also using MS EXCEL spread sheet and any statistical software such as SPSS, STATA etc. whichever convenient).	1
6	Computation measures of dispersion (ungrouped and grouped data) and computation of coefficient of variation. (Also using MS EXCEL spread sheet and any statistical software such as SPSS, STATA etc. whichever convenient)	1

S. No.	Title of the practical problems	No. of practical problems
7	Computation of raw and central moments	1
8	Measures of skewnwss and kurtosis using method of moments	1
9	Measures of Skewnwss using Box and whisker plot. (Also using MS EXCEL spread sheet and any statistical software such as SPSS, STATA etc. whichever convenient).	1
10	Scatter diagram, correlation coefficient (ungrouped data) and interpretation. Compute manually and check with computer output.	1
11	Fitting of lines of regression (Results to be verified with computer output)	1
12	Fitting of lines of regression and computation of correlation coefficient, Mean residual sum of squares, residual plot. (Also using MS EXCEL spread sheet and any statistical software such as SPSS, STATA etc. whichever convenient)	1
13	Combinatorial analysis and combinatorial probability	2
14	Geometrical probability	1
15	Conditional probability and Bayes theorem including sensitivity, specificity, predictive value positive and predictive value negative	3
16	Functions of random variables with linear and non linear transformations	2
17	Obtaining descriptive statistics of probability distribution	2
18	Fitting probability distributions in real data (Binomial, Poisson and Normal)	3
	Total number of practical problems	**25**