Course Title: Statistics and Probability **Credit: 3**
Course No: CSIT.216 **Number of period per week: 3+3**
Nature of the Course: Theory+ Lab **Total hours: 45**
Year: Second
Level: B. Sc. CSIT

1. Course Introduction

This course covers concept of descriptive statistics, probability, probability distributions, inferential statistics and their applications.

2. Objectives

At the end of this course the students should be able to:

- Know basic concepts of descriptive statistics, probability and their distributions, and inferential statistics and their applications in different areas.
- Identify existing pattern of data and their applications.
- Apply statistical tools and techniques in rational ways.
- Analyze the data scientifically and interpret them meaningfully

3. Specific Objectives and Contents

Specific Objectives	Contents
• Define statistics and probability, and state the scope, importance and limitations of statistics. • Explain the relations between statistics and information technology, and develop the concept of commuter software in association with statistics.	**Unit I: Concepts of Statistics and Probability (2 hr)** 1.1. Definition, importance, scope and limitations of statistics 1.2. Role of probability theory in statistics 1.3. Relations of statistics with information technology and e-methods.
• Define scales, attributes, variables and types of data, and also state the meaning of finite and infinite population, and sample, and distinguish between random and non-random sampling, • To organize the data, classify and tabulate them for presentation, and use appropriate diagrams & graphs for data presentation.	**Unit II: Concept of Population, Sample, Data and Variables and their types (3 hrs)** 2.1 Concept of attributes, scales, variables and their types, types of data, finite and infinite population, notation of sample, random and non-random sample. 2.2 Presentation of data- organization, classification and tabulation of data, rules of tabulation (strugs rule), diagrams and graphs. 2.3 Computational problems and examples
• Compute mean, median, mode, harmonic and geometric mean and partition values and interpret the results, and also state the properties • Compute absolute and relative variation, range, quartile deviation, standard deviation, mean deviation and coefficient of variation, Lorenz	**Unit III: Measures of Descriptive Statistics (8 hrs)** 3.1 Measures of locations- mean, median, mode, harmonic and geometric mean, partition values, and their use and properties. 3.2 Measures of dispersion- variation (absolute and relative), range, quartile deviation, mean deviation, standard deviation, coefficient of variation, Lorenz curve and gini-coefficient and

curve, gini-coefficient and also to interpret the result. • Describe the concept and use of skewness and kurtosis (by using partition values, central and raw moments).	their interpretations and use, 3.3 Measures of skewness and kurtosis, and their use. 3.4 Computational problems and examples
• To understand the terminologies of sample space, events, random experiment, trial, mutually exclusive events, equally likely cases, and to test the independence of the random variables. • To explain classical, statistical, axiomatic definitions of probability, basic principles of counting, permutation and combinations and compute them. • State additive, multiplicative, and conditional probability and compute probabilities, and state Bayes theorem and compute probability using Bayes theorem • Understand discrete & continuous random variables and to calculate probability distribution of a random variables • Compute expected values of discrete & continuous random variables	**Unit IV: Basic Probability Theory (5 hrs)** 4.1 Basic terminology in probability- sample space, events, random experiment, trial, mutually exclusive events, equally likely cases, favourable events, independent and dependent events, 4.2 Definition of probability- classical, statistical, subjective and axiomatic definitions, basic principles of counting, permutation and combinations, 4.3 Laws of probability- additive, multiplicative, and conditional probability, Bayes theorem with examples. 4.4 Random variables- discrete and continuous random variables, probability distribution of random variables 4.5 Expectation- expected value of discrete and continuous random variables, and mean and variance of random variable with illustrative examples. 4.6 Computational problems and examples
• To understand the marginal and joint probability distribution functions, mass and density functions, • Compute mean, variance, co-variance and correlation of random variables. • To know the independent & dependent random variables, • To know Bernoulli, binomial and Poisson random variables, and their distributions and moments, and also to compute their probabilities, test the normality of the distributions by using chi-square test. • Fitting binomial and Poisson distributions, • State the normal distribution and its moments, standardization of normally distributed random variable,	**Unit V: Probability Distributions (12hrs)** 5.1 Marginal and joint probability distributions, joint probability distribution of two random variables, marginal and joint probability mass functions and density functions 5.2 Mean, variance, co-variance, and correlation of random variables, independence of random variables 5.3 Discrete probability distributions- Bernoulli and binomial random variable and their distributions and moments. 5.4 Computing binomial probabilities and fitting binomial distribution (relate with chi-square test of the distribution pattern of the frequency). 5.5 Poisson random variable and its distribution and moments, and computing Poisson probabilities, and also fitting of Poisson distribution (relate with chi-square test of the frequency distribution). 5.6 Continuous probability distribution- normal

• To compute the areas under the normal curve, • Explain the negative exponential distribution and its moments, and also compute the probability.	distribution and its moments, standardization of normally distributed random variable, measurement of areas under the normal curve, 5.7 Negative exponential distribution and its moments, 5.8 Present the areas of application of above probability distributions. 5.9 Computational problems and examples
• To understand the definitions of chi-square, t and F random variables and their distributions and use them • Find the joint distribution of mean and sample variance of normal distribution	**Unit VI: Distribution of Chi-square, t and F (2 hrs)** 6.1 Definitions and properties of chi-square, t and F distribution and their random variables and their distributions and their comparisons 6.2 Find the mean and variance of these distribution (Proof is not required). 6.3 Computational problems and examples
• Understand simple random sampling methods and use it • Explain the sampling distribution and standard error and compute standard error and interpret the result • To know the distinction of descriptive and inferential statistics, point and interval estimation, • To understand the criteria of good estimator, maximum likelihood method of estimation • To estimate mean and variance in normal distribution, estimate the proportion in binomial distribution, • Compute the confidence interval of mean in normal distribution. • To know the step of testing of hypothesis, level of significance, types of error and power of the test. • Testing the hypothesis about mean in normal distribution in case of known variance (z-test) and unknown variance (t-test). • To carry out the ANOVA and also compute ANOVA table for one & two way classifications.	**Unit VII: Inferential Statistics (8 hrs)** 7.1 Concept of sampling its types (probability and non probability) with merits and demerits. 7.2 Steps of sample selection, determination of sample size. 7.3 Sampling distributions and standard error in both case (with and without replacement) 7.4 Distinction between descriptive and inferential statistics. 7.5 Concept of point and interval estimation, and criteria of good estimator, 7.6 Maximum likelihood method of estimation, and estimation of mean and variance in normal distribution, 7.7 Estimation of proportion in binomial distribution and confidence interval of mean in normal distribution 7.8 Concept of testing of hypothesis, level of significance, types of errors, power of the test, testing of hypothesis, concerning mean of a normal distribution in case of known variance and unknown variance. 7.9 Concept of analysis of variance (ANOVA), computation of one way and two way analysis of variance. 7.10 Computational problems and examples
• To understand and use correlation and regression in information technology • Compute correlation and regression coefficients and interpret the results,	**Unit VIII: Correlation and Regression (5 hrs.)** 8.1. Simple correlation- scatter diagram, Karl Pearson's correlation coefficient, and its properties, standard error, probable error, significant test of correlation coefficient.

and also state the properties. • Explain the assumptions of model, least-square estimators technique, and test of significance, and to compute the coefficient of determination and interpret the results. Use the analysis of variance in regression.	8.2. Computation of partial and multiple correlations and their consistency (up to three variables) 8.3. Simple linear regression- model and assumptions of simple linear regression, least square estimators of regression coefficients, standard error of estimate, test of significance of regression coefficients, coefficient of determination, and analysis of variance (up to three variables) 8.4. Computational problems and examples

Note: The figures in the parentheses indicate the approximate periods for the respective units. In addition to teaching hours (45), there will be 3 hours for reviews and discussions.

Evaluation System

Undergraduate Programs				
External Evaluation	**Marks**	**Internal Evaluation**	**Weightage**	**Marks**
End semester examination	60	Assignments & Lab	10%	40
(Details are given in the separate table at the end)		Quizzes	10%	
		Attendance	10%	
		Presentation	10%	
		Internal exams	50%	
		Group work	10%	
Total External	60	Total Internal	100%	40
Full Marks 60+40 = 100				

External evaluation

End semester examination: It is a written examination at the end of the semester. The questions will be asked covering all the units of the course. The question model, full marks, time and others will be as per the following grid.

Full Marks: 100, Pass Marks: 50, Time: 3 Hrs

Nature of question	**Total questions to be asked**	**Total questions to be answered**	**Total marks**	**Weightage**	**External exam marks**
Group A: Multiple choice*	20	20	20×1 = 20	20%	12
Group B: Short answer type questions	7	6	6×8 = 48	40%	24
Group C:Long answer type questions	3	2	2×16 =32	40%	24
			100	100%	60

.

Each student must secure at least 50% marks in internal evaluation in order to appear in the end semester examination. Failing to get such score will be given NOT QUILIFIED (NQ) and the student will not be eligible to appear in the end semester examinations.

Practical examination: Practical examination will be taken at the end of the semester. Students must demonstrate the knowledge and skill of the subject matter.

Internal evaluation

Assignment: Each student must submit the assignment individually. The stipulated time for submission of the assignment will be seriously taken.

Quizzes: Unannounced and announced quizzes/tests will be taken by the respective subject teachers. Such quizzes/tests will be conducted twice per semester. The students will be evaluated accordingly.

Attendance in class: Students should regularly attend and participate in class discussion. Eighty percent class attendance is mandatory for the students to enable them to appear in the end semester examination. Below 80% attendance in the class will signify NOT QUALIFIED (NQ) to attend the end semester examination.

Presentation: Students will be divided into groups and each group will be provided with a topic for presentation. It will be evaluated individually as well as group-wise. Individual students have to make presentations on the given topics.

Term paper: Term paper must be prepared by using computer in a standard format of technical writing and must contain the required number of pages. It should be prepared and submitted individually. The stipulated time for submission of the paper will be seriously taken as one of the major criteria of the evaluation.

Mid-term examination: It is a written examination and the questions will be asked covering all the topics in the session of the course.

Discussion and participation: Students will be evaluated on the basis of their active participation in the classroom discussions.

Instructional Techniques: All topics are discussed with emphasis on real-world application. List of instructional techniques is as follows:

- Lecture and discussion
- Group as well as individual work
- Self study and assignments
- Presentation by students
- Term paper writing
- Quizzes and guest lecture

Students are advised to attend all the classes and complete all the assignments within the specified time period. If a student does not attend the class(es), it is his/her sole responsibility to cover the topic(s) taught during that period. If a student fails to attend a formal exam/quiz/test, there won't be any provision for re-exam. Unless and until the student clears one semester he/she will not be allowed to study in the following semesters.

Prescribed Text

- Sheldon M. Ross. *Introduction to Probability and Statistics for Engineers and Scientists*, 3rd Edition, India, Academic Press, 2005.
- Shrestha, H.B. *Statistics and Probability- Concepts and Techniques*, EKTA Books Publication, Pvt. Ltd., reprint, 2008.

References

- Richard A. Johnson, Miller and Freunds. *Probability and Statistics for Engineers*, 6th Edition, Indian reprint, Pearson Education, 2001.
- Ronald E. Walole, R.H. Myers, S.L. Myers, and K. Ye. *Probability and Statistics for Engineers and Scientists*, 8th Edition, Indian reprint, Pearson Education, 2001.
- Aryal, T.R. *Fundamental Statistics- Concepts and Practices*, Viddharthee Publication, Pvt. Ltd., 2010.
- Martin, A. *Research Methods, Statistics, IT and e-Methods*. Icon Publication Pvt. Ltd, 2004.
- Yamane, T. *Mathematics for Economics*. Prentice-Hall of India Pvt. Ltd, 2000.
- Aryal, T.R. *Biostatistics-For Biology, Medical and Health Sciences*, Pinnacle Publication, Pvt. Ltd., 2011.
- Harry Frank & Steven C. Althoen. *Statistics Concepts and Applications*. Cambridge University Press (Low price edition), 1995.
- Murray R. Spiegel & Larry J. Stephens. *Statistics (Schaum's outlines)*, Tata McGraw-Hill Publishing Company Ltd, New Delhi, India, 2000.
- Kapoor J. N. and H.C. Saxena. *Mathematical Statistics*, S. Chand & Company Ltd., New Delhi, India, 2001.
- Gupta S. C. and Kapoor V. K. *Fundamentals of Mathematical Statistics*, Sultan Chand and Sons, 2007.
- Rohatgi V. K. and Ehsanes Saleh, A. K. MD. *An Introduction to Probability and Statistics*, John Wiley & Sons, 2005.
- Hoel, Port and Stone. *Introduction to Probability Theory*, Houghton Mifflin Company Boston, 1971.
- Hogg R.V and Criag, A.T. *Introduction to mathematical statistics, 3rd edition,* Academic Press, USA.
- Sukubhattu, N. P. *Probability Theory and Statistical Methods, 2nd edition,* Asmita Publications, Kathmandu, 2063BS.
- Miller and Fruend. *Modern Elementary Statistics*, Pearson Publication, 2007.
- Shrestha, Ganga. *Fundamental of Statistics*. ASAN Publications, Kathmandu,Nepal, 2006
- Feller, W. *An Introduction to Probability Theory and its Applications*, Vol. 1, Third edition, John Wiley and Sons, Singapore, 2000.
- Hoel, Port and Stone. *Introduction to Probability Theory*, Houghton Mifflin Company Boston, 1971.
- Mayer, P. L. *Introductory Probability and Statistical Applications*, second edition, Oxford and IBH Publishing Co. Pvt Ltd, New Delhi, 1970.
- Spiegel, M.R. *Theory and Problems of Statistics*, McGraw Hill Book Company, Singapore, 1992.

Note-

(i) Theory and practice should go side by side.
(ii) At least Excel and SPSS software should be used for data analysis.
(iii) It is recommended 45 hours for lectures and 15 additional hours for tutorial class for the completion of the course in the semester.
(iv) Home works and assignments covering the lecture materials will be given throughout the semester.

Specific objectives and contents of the practical problems

Specific objectives	Contents of the practical problems
• To organize and arrange raw data in appropriate classifications and tabulations for presentation and interpretation and use appropriate diagrams and graphs.	1. Arrange the data using strugs rule and present possible diagrams and graphs
• To compute mean, median, mode, harmonic mean and geometric mean and partition values, range, quartile,, standard deviation, mean deviation and coefficient of variation, Lorenz	2. Compute mean, median, mode, harmonic mean, geometric mean, partition values, range, quartile deviation, mean deviation, standard deviation, and coefficient of variation, Lorenz curve and gini-

curve, gini-coefficient, skewness and kurtosis using real data sets.	coefficient, skewness and kurtosis
• To compute probability distribution table of uni-variate and bi-variate data, and also calculate mean and variance using expectation	3. Develop probability distribution table of uni-variate data and bi-variate data, and compute mean and variance using expectation.
• To calculate marginal and joint probability table and mean, variance, co-variance and correlations, and test of independence of random variables, • To compute mean and variance of binomial and Poisson random variables, to test normality of binomial distribution using chi-square test, and also calculate areas under normal curve.	4. Calculate the marginal and joint probability distributions table and mean, variance, co-variance, and correlations of random variables, and test the independence of random variables, 5. Compute mean & variance of binomial and Poisson random variables; test the normality of binomial distribution using chi-square test, and compute the areas using normal curve.
• To compute sample size, sampling distributions and standard error, interval estimation of mean and proportion • To calculate t and z-test and one way and two analyses using real data sets.	6. Compute the sample size, sampling distributions and standard error with and without replacement, and also to compute interval estimation of mean and proportion 7. Compute t and z-test using real data, and one way and two way analysis of variance using real data sets.
• To compute simple, partial and multiple correlations, probable error, significant test of correlation coefficient using real data sets. • To fit linear regression and compute standard error of estimate, test of significance of regression coefficients, and coefficient of determination using real data sets.	8. Compute simple, partial and multiple correlations and derive probable error, significant test of correlation coefficient. 9. Fit simple linear regression, compute standard error of estimate, test of significance of regression coefficients, and coefficient of determination.

Note:

- Student must perform 3 hours of practical computer lab work every week.
- Students will develop the skills and knowledge on the calculations by using real data sets manual or computer software package.
- At least a problem is to be performed by each and every unit of the section of the above contents. Additional problems may be added subject to availability of time and skills of the students.
- The practical exam will be graded on the basis of the following marking scheme:

In-Semester Evaluation	20 %
Final Exam Written	60 %
Final Exam Oral	20 %

- The problems for practical computation are to be provided by respective teachers.