1. Introduction

Central Department of Statistics has decided to modify existing M. Sc. / M.A. semester course in Statistics after completion of one cycle of its running since 2012. From the practice it is realized that the semester system is relatively better than the annual system if it is implemented genuinely. It is also realized from the experience that the success of the semester system depends upon not only on the ways of evaluation but also on the methods of teaching and learning, active students participation in learning, evaluation and examination systems and overall designing of the syllabus.

When Central Department of Statistics started the semester system in M. Sc. / M.A. Statistics course, there were only few departments adopting the system. However, at present all of the central departments in university campuses have adopted the semester system and Tribhuvan University is trying to make uniformity regarding various factors within the system. Therefore, Central Department of Statistics has decided to make revision of the semester course of M. Sc. / M.A. in Statistics mainly to maintain uniformity among departments as regards the allocation of total credits for the program and evaluation scheme and also keeping in view of the international norms of the semester system.

The course contents of the program has been re-designed, developed and revised by the participation of experts, former Heads of the Department and faculty members of the Department. It consists of core theoretical papers, optional papers, computer based statistical computing (practical) papers and dissertation.

The re-structured syllabus for the **M. Sc. / M.A in Statistics in Semester System** with **major changes** is forwarded to the Dean's Office, IOST, TU for its approval.

2. Enrolment Quota

Considering the available laboratory space and its holding capacity, the Department has decided the quota of the number of students for enrolment in each batch. A maximum of forty students will be admitted in M. Sc. / M. A. in each batch of semester system. The admission of the students will be made on the merit basis through entrance examination and as per the rules and regulations of TU.

3. Eligibility for Student Intake

Students for the semester system will be enrolled based upon the rules and regulations laid down by TU. Students having a Bachelor Degree in Statistics or equivalent degree recognized by TU will be eligible to apply for admission in Semester System M. Sc. / M. A. in Statistics program.
Each applicant must appear and pass entrance examination conducted by the Central Department of Statistics, IOST. Admission will be made in the merit basis of the examination and marks obtained in Bachelor degree. Applicants with below cut-off point can be disqualified from enrollment.

4. Working Days and Class Duration

A. Working Days

Total number of working days for a semester will be 96 days equivalent to 16 weeks.

B. Class Duration

- Theory Paper

One credit will be equivalent to 16 teaching hours for theoretical papers in each semester. Consequently, theory papers of three credits will have three lecture hours per week.

- Statistical Computing (Practical) Paper

    1 practical class hour = 3 theory hours
    22 practical class hour = 4 credits

A computing paper of 4 credits will have 21-23 practical days in each semester. For Statistical Computing classes, groups of students will be formed having no more than 20 students in each group if the number of students exceeds 20. Three teachers (and / or instructors) will be allotted to instruct students simultaneously in each computing classes of each semester.

- **Dissertation and its Alternatives**

Dissertation is not mandatory to all students. The department will provide rules and criteria for the selection of students for dissertation. It will be of 4 credits and is allocated in fourth semester. The department will provide its orientation within the first two weeks from the date of the commencement of the fourth semester. The duration of the dissertation work will be of 96 working days after the completion of orientation. A single student will be facilitated by one supervisor with 1 hour of assistance/ supervision per week. Provision of the co-supervisor is also applicable, if necessary.

Regarding an alternative option to the dissertation work, students will have to take two theory papers of 2 credits each. The details of which are given in the syllabus.

C. Completion of Dissertation

Dissertation should be completed within 3 months after the end of the fourth semester to be regarded as a regular student.

5. Course Structure

The program is divided into four semesters (six months per semester) with a total duration of 2 years. The program contains 11 core courses, 5 optional courses, 3 computing (practical) courses, dissertation and two additional courses as an alternative option for dissertation. The distribution of courses in different semesters is shown below.

Distribution of Courses

Semester	Nature of the Course					Total Courses		Total Credit	Total Marks
	Core (Theory	Optional (Theory)	Statistical Computing (Practical)	Dissertation or Alternative Papers		Option A	Option B		
				Dissertation (Option A)	Alternative Paper (Option B)				
1st	6	0	1	0	0	7	7	22	550
2nd	3	2	1	0	0	6	6	19	475
3rd	2	3	1	0	0	6	6	19	475
4th	0	0	0	1	2	1	2	4	100
Total Papers	11	5	3	1	2	20	21		
Total Credit	33	15	12	4				64	
Total Marks	825	375	300	100					1600

The details of the program as regards to the distribution of subjects, credit hours, and marks in different semesters are shown in the following table. In total there are 64 credits (48 credits for theoretical papers, 12 credits for computing papers, and 4 credits for dissertation or additional exam based papers) with 1600 marks in total allocated in the program.

Course Structure

SN	Semester	Code	Subject	Credit	Marks
1	I	STA511	Mathematics for Statistics	3	75
2		STA512	Probability	3	75
3		STA513	Statistical Inference	3	75
4		STA514	Multivariate Analysis	3	75
5		STA515	Stochastic Processes	3	75
6		STA516	Programming Language	3	75
7		STA517	Statistical Computing-I (Practical)	4	100
		Total		22	550
8	II	STA521	Mathematical Demography	3	75
9		STA522	Sampling Theory	3	75
10		STA523	Design of Experiments	3	75
			Optional Paper (Any Two)	3×2 = 6	150
11		STA524	Econometrics		
12		STA525	Quality Control and Reliability		
13		STA526	Nonparametric Statistics		
14		STA527	Population Statistics		
15		STA528	Statistical Computing-II (Practical)	4	100
		Total		19	475
16	III	STA631	Bayesian Inference	3	75
17		STA632	Research Methodology	3	75
			Optional Paper (Any Three)	3×3 = 9	225
18		STA633	Biostatistics		
19		STA634	Environmetrics		
20		STA635	Time Series Analysis		
21		STA636	Operations Research		
22		STA637	Survival Analysis		
23		STA638	Actuarial Statistics		
24		STA639	Statistical Computing-III (Practical)	4	100
		Total		19	475
	IV	Any <u>One</u> between Dissertation and Two Alternative Papers			100
25		STA641	Dissertation	4	
		Alternative Papers (2 Credits each)			
26		STA642	Meta Analysis	2	
27		STA643	Nonparametric and Categorical Data Modeling	2	
		Total		4	100
			Grand Total	**64**	**1600**

6. Course Details

Semester I

Course Title: Mathematics for Statistics **Full Marks: 75**
Course Code: STA511 **Pass Marks: 37.5**
Total Credits: 3
Total Lecture Hours: 48

Course Objective: This course enables students to acquire knowledge of the mathematics needed for the study of advanced theory of Statistics. Students will also develop their competence in applying mathematical techniques in solving problems in Statistics.

UNIT 1: Numerical Analysis 15 hrs

Algebraic and Transcendental Equations: Bisection method, Iteration method, Newton Raphson method.

Interpolation and Extrapolation: Finite differences (forward, backward and central), Newton's formula for interpolation (forward, backward), Lagrange's interpolation formula.
Numerical Differentiation and Integration: General formula for numerical differentiation and integration, Trapezoidal rule, Simpson's 1/3 and 3/8 rules.

Ordinary Differential Equations**:** Solution of Taylor's series, Picard's method, Euler's method, Runge-Kutta method.

UNIT 2: Real Analysis 30 hrs

Sequences and Series: Sequences and series of functions, Point wise and uniform convergence, Cauchy general principle of convergence for sequence, Limit superior and limit inferior.
Power series: Radius of convergence, Convergence of power series.
Fourier series: Periodic function and its properties, Sum of Fourier series, Fourier series of even and odd functions.

Integration**:** Review of Riemann integral, Riemann-Stieltjes integral, Condition of integrability, Mean value theorem, Integration by parts, Improper Integral, Convergence of improper integrals, Convergence of Beta and Gamma function.

Function of several variables: Definition of Limit &Continuity, Partial Derivative, Euler's Theorem, Jacobian, Maxima and Minima, Multiple integral, Dirichlet's theorem, Liouville's expansion to Dirichlet's theorem, Parametric integration

Review of the overall course 3 hrs

Reference Books:

1. **Apostol, T.M.** (2002): *Mathematical Analysis*, Narosa Publishing House, New Delhi
2. **Bartlett, R.G. and Sherbet, D.R**. (1994): *Introduction to Real Analysis*, John Wiley and Sons, New York
3. **Chatterjee, D.** (2005): *Real Analysis*, Prentice-Hall of India, India
4. **Malik, S.C. and Arora, S.** (1992): *Mathematical Analysis*, New Age International, India
5. **Sastri, S.S.** (2003): *Introductory Methods of Numerical Analysis*, Prentice-Hall of India, India

Course Title: Probability **Full Marks: 75**
Course Code: STA512 **Pass Marks: 37.5**
Total Credits: 3
Total Lecture Hours: 48

Course Objective: To impart knowledge and improve level of understanding of probability theories and probability distributions along with their applications.

UNIT 1: Sets and Fields 10 hrs

Limit and field: Event, algebra of sets, limit of sequence of sets, limit superior and limit inferior, field, □ field, minimal field, monotone field, Borel field, ring. Function, measure and random variable: definition of function, set function, inverse function, measure, measure space and measurable function, probability measure, random variable.

UNIT 2: Probability Space and Functions 13 hrs

Probability Space: Axiomatic definition of probability, probability function and its properties, probability space, discrete, finite, countable and general probability spaces with examples.
Conditional Probability: Conditional probability measure and independence of events, occupancy problems.
Distribution Function: Definition and its properties, distribution of vector random variables, transformation of random variables. Expectation: Definition, expectation in univariate and multivariate distributions and independent random variables, conditional expectation, expectation of linear combinations, relation between expectation and cumulative distribution function.
Characteristic Function: Characteristic function and its properties, inversion formula and uniqueness theorem, examples of use of inversion formula

UNIT 3: Probability Distributions 22 hrs

Multinomial Distribution: Probability mass function, moment generating and characteristic function, moments, covariance and correlation, distribution fitting and examples
Extreme Value Distributions: Probability density and distribution functions, moments, properties and examples.
Distribution of Order Statistics: Distribution of kth order statistics, joint and marginal distributions of order statistics, problems and examples.
Generalized Power Series Distribution: Unified PMF, its special cases (binomial, Poisson, negative binomial).
Prior and Posterior Distributions: Meaning and examples including cases where binomial, beta, exponential, gamma, Poisson, negative binomial distributions are involved.
Compound Negative Exponential Distribution: Compounding of distributions, its moments.
Mixed Type Distribution: Mixed random variable, meaning and examples, computation of moments of mixed random variables

Review of the overall course 3 hrs

Reference Books:

1. Bhat, B.R. (1999): Modern Probability Theory - An Introductory Textbook, New Age International, New Delhi
2. Biswas, S. (1991): Topics in Statistical Methodology, Wiley Eastern, India
3. Rohatgi, V.K. and Saleh, A.K.Md.E. (2005): An Introduction to Probability and Statistics, John Wiley and Sons, Singapore
4. Hogg, R.V. and Tanis, E.A. (2001): Probability and Statistical Inference, Pearson Education, India
5. Meyer, P.L. (1970): Introductory Probability and Statistical Applications, Addison-Wesley, USA.
6. Shrestha, S. L. (2011) Probability and Probability Distributions, S. Shrestha, Kathmandu.
7. Chandra, T.K. and Chatterjee, D. (2003): A First Course in Probability, Narosa Publishing House, India
8. Hoel, P.G., Port, S.C. and Stone, C.J. (1971): Introduction to Probability Theory, Universal Book Stall, New Delhi.

Course title: Statistical Inference **Full Marks: 75**
Course Code: STA513 **Pass Marks: 37.5**
Total Credit Hour: 3
Total Lecture Hours: 48

Course Objective: The objective of this course is to impart the knowledge of inferential statistics in decision- making process.

UNIT 1: Estimation 10 hrs

Minimal sufficiency, likelihood equivalence, completeness, uniformly minimum variance unbiased estimator (UMVUE). Fisher information. Lower bound to variance of estimators, necessary and sufficient condition for minimum variance unbiased estimator.
Method of estimation- maximum likelihood method, method of moments, method of minimum Chi-square, method of least square. Asymptotic properties of maximum likelihood estimator.

UNIT 2: Interval Estimation 13 hrs

Construction of shortest length confidence interval, Uniformly Most Accurate Unbiased confidence interval, construction of confidence interval for population proportion (small and large samples) and between two population proportion, confidence interval for mean variance of a normal population, difference between mean and ratio of two normal population.

UNIT 3: Testing of Hypothesis 14 hrs

General concept on simple and composite hypothesis, two types of errors, level of significance, power and size of a test. Most powerful test – Neymann Pearson's lemma and its application. Uniformly most powerful test-application to standard statistical distribution, unbiased test. Likelihood ratio test- Principle and properties, derivation of likelihood ratio test for testing means and variance in exponential families.

UNIT 4: Sequential Tests 8 hrs

Sequential probability ratio test(SPRT), Derivation of SPRT for testing parameter of binomial, exponential and Poisson distribution. Operating characteristic function. Average sample number.

Review of the overall course 3 hrs

Reference Books:

1. Rohatgi, V.K. and Saleh, A.K. Md.E. (2005) An Introduction to Probability and Statistics, Second Edition, John Wiley.
2. Kale, B.K. (1999), A First Course on Parametric Inference, Narosa Publishing House.
3. Lehmann E.L. (1986), Theory of Point Estimation, John Wiley and Sons.
4. Lehmann E.L. (1986), Testing Statistical Hypotheses, John Wiley and Sons.
5. Zacks,S. (1971), Theory of Statistical Inference, John Wiley and Sons.

Course Title: Multivariate Analysis **Full Marks: 75**

Course Code: STA514 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: This course has two-fold objectives. First objective is to provide fundamental knowledge of multivariate normal distribution and multivariate statistical methods with their applications. The second objective is to impart the theoretical knowledge of advanced statistical methods with their applications based on computer.

UNIT 1: Multivariate Normal Distribution (MVND) 7 hrs

Density function and characteristic function of MVND, Distribution of linearly transformed multivariate normal random vector, Marginal and conditional distribution of MVND, Necessary and sufficient condition for independence in MVND

UNIT 2: Estimation of Multivariate Normal Parameters 6 hrs

Concept about Sampling from MVND, MLEs of mean vector and dispersion matrix (derivation not required), Properties and distributions of MLEs, Wishart distribution (derivation not required) and its properties, MLEs of simple, partial and multiple correlation coefficient and their distributions (derivation not required)

UNIT 3: Hypothesis Testing in Multivariate Normal 7 hrs

Hotelling's T^2 statistic as a generalization of square of Student's statistic, Defining Hotelling's T^2 statistic from likelihood ratio test, Distribution of Hotelling's T^2 statistic and its invariance property, Applications of T^2 statistic in hypothesis testing (one sample and two sample problems), Distance between two populations, Mahalanobis' D^2 statistic

UNIT 4: Principal Component Analysis 6 hrs

Model formulation, number of components and component structure, extraction of principal components, maximum likelihood estimators of principal components and their variances

UNIT 5: Factor Analysis 7 hrs

Factor Analysis based on Principle Axis Factoring Approach and Principle Components Approach, **o**rthogonal factor model, oblique factor model, estimation of factor loadings, communalities, rotation of factors, factor scores and their applications, maximum likelihood estimators for random orthogonal factors, tests of hypothesis in factor models

UNIT 6: Discrimination and Classification Analysis 7 hrs

Separation and classification for two populations, classification with two multivariate normal populations, Fisher's Discriminant function, classification with several populations

UNIT 7: Multivariate analysis of variance 5 hrs

Multivariate One-Way Analysis of Variance Model (MANOVA), Wilks test, Roy's Test

Review of the overall course 3 hrs

Reference Books:

1. **Anderson, T. W.** (1983): *An Introduction to Multivariate Statistical Analysis,* 3rd edition, John Wiley and Sons.
2. **Rao, C. R.** (2002): *Linear Statistical Inference and its Applications,* John Wiley and Sons
3. **Johnson, R. A**. and **Wichern, D. W.** (2006): *Applied Multivariate Statistical Analysis*, 5th edition, Prentice Hall of India
4. **Hardle, W and Simar L**. (2007): *Applied Multivariate Statistical Analysis,* 2nd Edition, Springer

Course Title: Stochastic Processes **Full Marks: 75**

Course Code: STA515 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: To impart knowledge on theory and practices of stochastic processes with their applications from physical sciences and engineering fields.

UNIT 1: Introduction to Stochastic Processes 12 hrs

Classification of stochastic processes according to state space and time domain, probability generating function and its properties, random walk, return to origin probability, gambler's ruin problem, Galton-Watson branching process, mean and variance of generation size, probability of ultimate extinction, total progeny.

UNIT 2: Markov Chain 9 hrs

Introduction, transition probability, absolute probability, Chapman-Kolmogorov equations, nth step tpm for two-state MC, spectral analysis of two-state MC, classification of states of MC, countable state MC.

UNIT 3: Discrete and Continuous Markov processes in continuous time 12 hrs

Poisson process& its properties, birth and death processes, diffusion process, Brownian motion process, Kolmogorov forward and backward diffusion equations, Martingale.

UNIT 4: Renewal Theory 6 hrs

Introduction, renewal function, integral equation of renewal theory, stopping time and Wald's equation, spent and residual time distribution, elementary renewal theorem.

UNIT 5: Theory of Queue 6 hrs

Introduction, operating characteristic of queue theory, M/M/1 and M/M/s queue system.

Review of the overall course 3 hrs

Reference Books:

1. Bhat, B. R. (2000). Stochastic Models- Analysis and Applications, New Age International Publishers.
2. Feller, William (1968). An Introduction to Probability Theory and its Applications, Vol. 1 (Third Edition.), John Wiley.
3. Medhi, J. (2009). Stochastic Processes, 3rd Edition, New Age International Publishers.
4. Karlin, S. and Tyalor, H.M. (1975). A First Course in Stochastic Processes, Second Edition. Academic Press.
5. Ross, Sheldon M. (1983). Stochastic Processes, 2nd Edition, John Wiley and Sons, Inc.
6. . Hoel, P.G., Port, S.C. and Stone, C.J. (1972). Introduction to Stochastic Processes, Houghton Miffin& Co.
7. Shrestha, H.B. (2009). Stochastic Processes, An Introductory Text, Ekta Books

Course Title: Programming Language **Full Marks: 75**
Course Number: STA516 **Pass Marks: 37.5**
Total Credits: 3
Total Lecture Hour: 48

Course Description: This course is designed to develop acquaintance with fundamental concepts of program design and computer programming. The course starts with the basic concepts and also includes the concepts of C programming including data types, operators, control statements, arrays, functions, pointers, structures, unions, data files, and numerical analysis

Course Objective: On completion of this course, students will be able to develop their knowledge in program design and computer programming and they will be able to develop small to medium size computer programs using different concepts of C programming language

UNIT 1: Introduction to programming languages 4 hrs
Evolution of programming languages, structured programming, the compilation process, object code, source code, executable code, operating systems, interpreters, linkers, loaders, fundamentals of algorithms, flow charts, Introduction to software development, Number System Representation

UNIT 2: C Language Fundamentals 12 hrs
Character set, Identifiers, Keywords, Data Types, Constant and Variables, Statements, Expressions, Operators, Precedence of operators, Input-output Assignments, Control structures, Decision making and Branching, Decision making & looping

UNIT 3: C Functions 4 hrs
User defined and standard functions, Formal and Actual arguments, Functions category, function prototypes, parameter passing, Call-by-value, Call-by-reference, Recursion, Storage Classes

UNIT 4: Arrays and Strings 5 hrs
One dimensional Array, Searching, Sorting, Multidimensional Array, Matrix operations, String Manipulation

UNIT 5: Pointers 6 hrs
Pointer variable and its importance, Pointer Declarations, Passing Pointers to a Functions, Pointers and One-dimensional Arrays, Dynamic Memory Allocation, Operations on Pointers, Pointers and Multi-dimensional Arrays, Arrays of Pointers, Pointer to pointer, Linked list

UNIT 6: Structures, Unions 5 hrs
Defining a Structure, Processing a Structure, User Defined Data Types (typedef), Structures and Pointers, Passing Structures to Functions, Self-referential Structures, Unions

UNIT 7: File Handling 4hrs
Why Files, Opening and Closing a Data File, Reading and Writing a Data File, Processing a Data File, Unformatted Data Files, Concept of Binary Files

UNIT 8: Numerical Analysis 5 hrs
Errors in Numerical Calculations, Roots of Algebraic and Transcendental Equations by Bisection and Newton-Raphson Methods, Lagrange interpolation, Least Square Straight Line Fitting

Review of the overall course 3 hrs

Reference Books

1. Programming in C - Gottfried Byron
2. The ‘C’ programming language - B.W.Kernighan, D.M.Ritchie
3. Introductory Methods of Numerical Analysis, S.S. Sastry, PHI
4. Programming in ANSI C - Balaguruswami
5. C The Complete Reference - H.Sohildt
6. Let us C - Y.Kanetkar
7. A Structured Programming Approach using C – B.A. Forouzan & R.F. Gillberg
8. Computer fundamentals and programming in C – Pradip Dey & Manas Ghosh

Course Title: Statistical Computing-I **Full Marks: 100**

(Practical Paper) **Pass Marks: 50**

Course Code: STA517 **Total Credits: 4**

Total Lectures:21-23(3 Hours / Lecture)

Total Duration: 64 hours

Course Objective: The objective of this course is to enable the students to apply theories learnt in solving statistical problems.

SN	Subject Area	No. of Practicals
1	Statistics for Mathematics	5
2	Probability Theory	5
3	Statistical Inference	5
4	Stochastic Processes	5
5	Multivariate Analysis	5
	Total	25

Semester II

Course Title: Mathematical Demography **Full Marks: 75**

Course Code: STA521 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Objective of the Course: To acquaint and make the students capable of measuring the various demography variables by using direct and indirect techniques.

UNIT 1: Age distribution 5 hrs

Sources of errors in age statement
Age curve
Age pyramid
Ageing Index
Age adjustment
- Method of parabola
- Adjustment of error in age group 0-4 years
- Adjustment of error in age group 5-9 years
- Adjustment of error in age group 75+ years

UNIT 2: Migration 11 hrs

Definition (4 hrs)
Causes of migration
Neo Marxists Theory
Migration in Nepal
Measuring migration

Direct method (3hrs)
- Based on Birth Place
- Based on Place of residence
- Based on Duration of Residence
- Based on Residence on a fixed prior data

Indirect Method (4hrs)
- Residue Method
- Survival Ratio
- Varying Survival Ratio method
- Net Migration of children
- National Growth Rate method
- Net Reproduction method
- Population Redistribution, width example of Nepal

UNIT 3: Marital rate 6 hrs

Risk for first marriage

Measuring mean age at marriage
- Hajnal' Method
- Cohort method
- Stable population method
- Vadelle Walle method
- Singh's method

UNIT 4: Nuptuality Models 3 hrs

Coale's Gompertz curve
Coale and MC Neil 's Extension of Gompertz Extension of the curve
T.James Tussel's Extension of Gompertz curve

UNIT 5: Fertility Models 5 hrs
- Parabolic
- Gomperz
- Coale- Tussel's model
- Brass model

UNIT 6: Indirect Techniques in measuring Fertility levels 8 hrs

Its needs
Nature of data on children ever born
Errors in fertility data
El-Brady correction
Coale- Demeny 's Method of Estimation of TFR
Estimation of Ever-fertile women
Estimation of number of women with known parity
Adjustment of ASFR, Brass' s P/F Method
Adjustment of ASFR, Coale- Tussel's model
Comparing Period fertility rates with a hypothetical cohort
- Cohort Parity increase method
- Ten years Survival method

UNIT 7: Indirect Techniques in measuring Mortality levels 7 hrs

Brass's Method of Estimation of Infant And Child Mortality
Ten years survivorship method for Estimation of birth and death rates from stable population
Singh's Method for Estimation of Birth and Death rates from Census data
Indirect method of Estimation of IMR
James, McCann's Method's of Estimation of life expectancy at Birth
Singh's method of Estimation of life expectancy at birth
Vig's Relation between life expectancy at birth and death rates

Review of the overall course 3 hrs

Reference Books:

1. O.P Vig (1976): India's Population (A study through extension of Stable Population Technique), Sterling Publisher PVT, LTD New Delhi
2. Singh M.L (1995): Some Measures of Demographic variables, Kathmandu
3. Singh M.L, Saymi,S.B (1997): An introduction to Mathematical Demography , Kathmandu
4. Singh M.L (2000): Population Growth and Migration, CDS, T.U

Course Title: Sampling Theory **Full Marks: 75**

Course Code: STA522 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: The objective of this course is to make students familiar with higher knowledge in survey sampling. After completion of this course, the students will be able to carry out survey sampling independently.

UNIT 1: Review of Some Preliminary Results

Review of same important results in simple random sampling with without replacement, stratified sampling, ratio and regression estimators. 2 hrs

UNIT 2: Stratified Sampling

Stratified sampling with varying probabilities of selection for estimation of population mean, Horvizt-Thompson estimator for population mean, Post stratification, Difference estimator in stratified sampling 8 hrs

UNIT 3: Ratio Estimator

Concept of Multivariate Ratio Estimator, Multivariate ratio estimator for two auxiliary variables, Comparison multivariate ratio estimators with customary ratio estimator having single auxiliary variable and with simple random sampling without replacement, Combined and separate ratio estimators, Product estimator 8 hrs

UNIT 4: Super Population

An outline of fixed and super population approaches, Model base estimates of parameter for simple random sampling and ratio estimate. 6 hrs

UNIT 5: Multiphase Sampling

Concept of multiphase sampling, Double sampling for stratification, Optimum allocation, Estimate of variances in double sampling for regression. 6 hrs

UNIT 6: Cluster Sampling

Cluster sampling variance estimate of inter-cluster correlation of equal cluster size, Cluster sampling for unequal size. 6 hrs

UNIT 7: Sub-Sampling

Concept of sub-sampling, Two-stage sampling, Equal First Stage Units: Estimation of the Population Mean, Unbiased estimate of sampling variance, Three- stages sampling with equal first and second stages units. 5 hrs

UNIT 8: Variance Estimation

Concept of variance estimation, Variance estimation in complex survey, interpenetrating sub-sampling method, Method of random groups: case of independent random group, The bootstrap. 4 hrs

Review of the overall course 3 hrs

References Books:

1. **Cochran, W. G.** (1977): *Sampling Techniques*, Wiley-Eastern, India
2. **Mukhopadhyay,P.** (1998): *Theory and Methods of Survey Sampling.* Printice Hall of India, India
3. **LohrS.L.** (1999): *Sampling: Design and Analysis.* Duxbury Press. USA.
4. **Raj D., Chandhok, P.** (1999): *Sample Survey Theory*, Narosa Publishing House,India
5. **Wolter, K.M.** (1985): *Introduction to Variance Estimation.* Sringer-Verlag,New York
6. **Chaudhuri, A.** (2010): *Essentials of survey sampling*, PHI Learning Pvt, India
7. **Sampath, S.** (2005): *Sampling Theory and Methods*, Narosa Punlishing House. India.

Course Title: Design of Experiments **Full Marks: 75**

Course Code: STA523 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: To impart knowledge and improve level of understanding of Experimental Designs along with their applications.

UNIT 1: Basic Designs 7 hrs

Review of linear estimation.

Linear Models: Fixed effects, random effects and mixed effects models, computation of sum of squares and construction of ANOVA table, analysis of covariance with a single covariate in randomized block design.

UNIT 2: Nested and Split Plot Designs 9 hrs

Nested Design: The two stage nested design, crossed and nested factors, design layout and model specification, estimation of model parameters, computation of sum of squares, construction of ANOVA table.

Split Plot Design: Design layout and model specification, least square estimation of parameters, computation of sum of squares, construction of ANOVA table.

UNIT 3: Incomplete Block Designs 8 hrs

Balanced Incomplete Block Design: Design layout and model specification, intra-block estimation, computation of sum of squares (adjusted and unadjusted), recovery of inter-block information.

Other Incomplete Block Designs: Connectedness, partially balanced incomplete design and Youden Squares.

UNIT 4: Two-level Factorial and Fractional Factorial Designs 15 hrs

2^k factorial design and its analysis

Confounding in 2^k Design: Confounding 2^k design in two, four and 2^p blocks, construction of blocks, partial and complete confounding.

Fractional Factorial Designs: Two-level fractional factorial design and its analysis, alias structure and design resolutions, one-half and one-quarter fractions of the 2^k design, the general 2^{k-p} fractional factorial design.

UNIT 5: Response Surface Methodology 5 hrs

The first and second order response surface models, method of steepest ascent, location of stationary point, fitting and analysis of response surfaces.

Review of the overall course 3 hrs

Reference Books:

1. **Montgomery, D. C. (2003):** *Design and Analysis of Experiments*, John Wiley and Sons, Singapore.
2. **Charles R. H. and Turner, K. V. Jr. (1999):** *Fundamental Concepts in the Design of Experiments*, Oxford University Press, New York.
3. **Cochran, W. G. and Cox, G. M. (1992):** *Experimental Designs, 2nd Edition*, John Wiley and Sons, Inc., USA.

Course Title: Econometrics **Full Marks: 75**

Course Code: STA524 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: To provide an elementary yet comprehensive introduction to econometric regression analysis including time series analysis.

UNIT 1: Multiple Linear Regression (Matrix Approach) 10 hrs

The *k* variable regression model, model specification, assumptions in matrix notation, OLS estimation of parameter vector, variance-covariance matrix of estimated parameter vector, standard error of estimates, properties of vector estimate, unadjusted and adjusted multiple coefficient of determinations, correlation matrix, hypothesis testing of regression coefficients, testing of goodness of fit by analysis of variance, prediction.

Problems and examples

UNIT 2: Violation of Assumptions 15 hrs

Multicollinearity: The nature of multicollinearity, estimation and detection of multicollinearity, consequences and remedial measures

Heteroscadasticity: The nature of heteroscadasticity, OLS estimation in the presence of heteroscadasticity, detection of heteroscadasticity: Goldfeld-Quandt test, Breusch-Pagan-Godfrey test, White's test, treatment of heteroscedasticity: method of weighted least squares, White's Heteroscedasticity - consistent variance and standard errors.

Autocorrelation and residual analysis: The nature of autocorrelation, detection and consequences of autocorrelation, Durbin Watson test, remedial measures: Changing the functional form, Cochrane-Orcutt iterative procedure.

UNIT 3: Binary Logistic Regression 5 hrs

Categorical dependent variable in regression, binary logistic regression: model specification, assumptions, estimation and interpretation of model parameters, odds ratio.

UNIT 4: Time Series Analysis 15 hrs

Time series and stochastic processes, stationary and non-stationary time series or process, detection of stationarity: series plot against time, test based upon correlogram, autocorrelation function (ACF), random walk, unit root problem, Dickey-Fuller test, transforming non-stationary time series.

Time series models: Autoregressive (AR) model, moving average (MA) model, autoregressive and moving average (ARMA) model, autoregressive integrated moving average (ARIMA) model, Box-Jenkins methodology, estimation of the ARIMA model and forecasting.

Co-integration: Definition, test for detection of cointegration, Durbin-Watson test for cointegration
Distributed lag models: Model specification, estimation: Koyck approach and Almon lag.

Review of the overall course 3 hrs

Reference Books:

1. **Gujarati, D.N. and Sangeetha** (2007): *Basic Econometrics*, Tata McGraw-Hill, New Delhi
2. **Ramanathan, B.** (2002): *Introductory Econometrics with Applications*, South-Western Thomson Learning, Singapore
3. **Maddala, G. S.** (2002): *Introduction to Econometrics*, John Wiley and Sons
4. **Drapper, N. R. and Smith, H. (1998)** *Applied Regression Analysis*, Third edition, Wiley, New York.

Course Title: Quality Control and Reliability **Full Marks: 75**

Course Code: STA525 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: The objective of this course is to impart knowledge of Statistical Process Control, Sampling Inspection and Reliability and to develop the skills of applying statistical techniques to industrial data.

UNIT 1: Statistical Basis of the Control chart: 21 hrs

- Introduction to Quality control; General theory and review of control charts
- Control chart for variables (x-bar, R and s charts) and attributes (p, np, c, u charts)
- Operation Characteristic, ARL of control charts; Economic design of x-bar chart
- CUSUM chart: i) Algorithm and ii) V mask procedure for monitoring process mean using CUSUM chart
- Modified control chart and acceptance control chart
- Process capability: Specification Limit and Tolerance Limit, Definition and uses of Process Capability Indices Cp, Cpk, and Cpm

UNIT 2: Sampling Inspection 12 hrs

- Sampling Inspection Plans:
  - Single, Double and Multiple sampling plans
  - Concept and Interpretation of LQL and AQL, consumer's risk and producer's risks.
  - OC function, construction of OC curves
- Method for estimating n and c using large sample
- Corrective Sampling Plan:
  - Rectifying Inspection Program
  - Interpretation of AOQ, AOQL, ATI, ASN,
  - Curtailed and semi-curtailed inspection plan
- Sampling plan by variables, Sequential and chain sampling plans

UNIT 3: Measures of Reliability 12 hrs

- Concept and measures of Reliability:
  - Failure time Distribution and Reliability function
  - Hazard rate and General equation of failure rate distribution; Mean time to failure
- Reliability measures from common life testing models:
  - Exponential, Weibull, Lognormal, Rayleigh and Bath-tub models
- Reliability of system
  - Series system, Parallel system and Series-parallel configuration
- Reliability of maintained system
  - Concepts and interpretation of Maintainability and Availability.
  - System availability, Preventive maintenance

Review of all the course 3 hrs

Reference Books:

1. Montgomery, D. C. (2004). Introduction to Statistical Quality Control, John Willey and Sons
2. Grant, E. L. and Leavenworth, R.S. (2004). Statistical quality Control, Tata McGraw Hill
3. Bisws, S. (1997) Statistical Quality Control , New Age India
4. Sinha, S. K. and B. K. Kale (1980). Life Testing and Reliability Estimation, Willey Eastern
5. Khatiwada, R. P. (2013). An Introduction to Statistical Quality Control and Reliability, Quest publication, Kathmandu, Nepal.

Course Title: Nonparametric Statistics **Full Marks: 75**

Course Code: STA526 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: This course deals with statistical inference when parametric distributions are not assumed and presents the theory and procedures of decision making in absence of rigid distributional assumptions.

UNIT 1: Order Statistics **10 hrs**

- Probability integral transformation
- Joint and marginal distribution of r^{th} order statistics
- Moments of order statistics
- Distribution of median and range
- Asymptotic distribution of order statistics
- Confidence interval estimates for population quintiles
- Hypothesis testing for population quintiles

UNIT 2: Distribution-Free Statistics **10 hrs**

- Distribution-free statistics over a class
- Counting statistics
- Ranking statistics
- U-statistics: one sample and two sample U-Statistics and their asymptotic properties
- Asymptotically distribution-free statistics

UNIT 3: Non-Parametric Tests **25 hrs**

- One sample tests
    - Binomial test
    - Tests based upon runs
        - Exact null distribution of R
        - Moments of the null distribution of R
        - Asymptotic null distribution of R
    - Sign test
    - Wilcoxon signed rank test.
- Tests of goodness of fit
    - Chi-square test
    - Kolmogorov-Smirnov test

- Two sample tests
    - Wald-Wolfowitz runs test
    - Kolmogorov-Smirnov two sample test
    - Median test
    - Mann-Whitney U test
- Several sample tests
    - Kruskal-Wallis one way ANOVA test
- Measures of association
    - Kendall's tau coefficient
    - Spearman's coefficient
    - Contingency coefficient
    - Coefficient of concordance
    - Friedman's two way analysis of variance by ranks

Review of the overall course 3 hrs

Reference Books:

1. **Gibbons, J.D.** (1985): *Nonparametric Statistical Inference*, Marcel Dekker, New York
2. **Randles, R.H. and Wolfe, D.A.** (1979): *Introduction to the Theory of Nonparametric Statistics*, John Wiley and Sons, New York
3. **Rohatgi, V.K. and Saleh, A.K.Md.E.** (2005): *An Introduction to Probability and Statistics*, John Wiley and Sons, New York
4. **Conover, W.J.** (1980): *Practical Nonparametric Statistics*, John Wiley and Sons, New York

Course Title: Population Statistics **Full Marks: 75**

Course Code: STA527 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Objectives: The main objective of this course is to familiarize students with population related statistics and also to broaden student's understanding about population dynamics of Nepal.

UNIT 1: Sources of Data 2 hrs

Censuses data, demographic health surveys data, other sample surveys carried out by different agencies, vital registration system of Nepal, Service statistics of Nepal
Examples with recent data sets of Nepal

UNIT 2: Population distribution of Nepal 4 hrs

Patterns, characteristics, trends and composition of population of Nepal (size, growth, density, age-sex composition, demographic, economic and caste/ethnicity, etc. composition of population), dependent, child and aging population of Nepal.
Examples with recent data sets of Nepal

UNIT 3: Characteristics, trends and composition of nuptiality of Nepal 6 hrs

Nuptiality patterns, levels and trends of age at marriage, marital characteristics of the population by eco/region, urban/rural, caste/ethnicity, etc., differentials and determinants of age at marriage of Nepal by different characteristics of population.
Examples with recent data sets of Nepal

UNIT 4: Characteristics, patterns, trends and composition of fertility of Nepal 9 hrs

Level, trends and patterns fertility of Nepal by different characteristics of population (eco/region, urban/rural, socio-economic, demographic, caste/ethnicity, etc), differentials and proximate determinants of fertility, level of wanted and unwanted fertility, family planning impact on fertility and abortion statistics.
Examples with recent data sets of Nepal

UNIT 5: Levels and trends of mortality and morbidity of Nepal 8 hrs

Level, trends and patterns of infant, child, adult and maternal mortality of Nepal, life expectancy, differential of high risk mortality of Nepal by different characteristics (eco/region, urban/rural, caste/ethnicity, etc). morbidity statistics, its levels and patters.
Examples with recent data sets of Nepal

UNIT 6: Migration of Nepal 7 hrs

Streams and trends of migration, patterns and differentials of migration at individual, household and village levels by socio-economic, cultural and demographic characteristics of migrants, reasons and causes of migration, levels of internal and international migration and characteristics of migrant population.
Examples with recent data sets of Nepal

UNIT 7: Levels and trends of population growth of Nepal 4 hrs

Levels, trends, differentials and composition of population growth, population projection by different methods, and causes of population growth of Nepal.

UNIT 8: Human Development Statistics 8 hrs

Concept and implication of Human Development Index (HDI): Measures, dimensions, indicators, constructing of HDI, Gender related development index (GDI): Measures, dimensions, indicators, constructing GDI, Gender empowerment measure (GEM): Measures, dimensions, indicators, constructing GEM, Human poverty index (HPI): Measures, dimensions, indicators, constructing HPI, Levels and patterns of human development of Nepal.

Review of the overall course 3 hrs

Reference Books:

- Aryal, T.R. (2011). Fertility Dynamics of Nepal, Ekta Book Distributors, Kathmandu.
- Aryal, T.R. (2010). Nuptiality, Gyankunja Prakashan, Kirtipur, Kathmandu.
- Aryal, T.R. (2011). Mortality of Nepal, Prime Publication, Teku, Kathmandu.
- Aryal, T.R. (2008). Migration and Occupational Mobility in Nepal, Paluwa Prakashan, Bagbazzar, Kathmandu.
- Ministry of Population and Health. (1996. 2001, 2006, 2011). Nepal Demographic and Health Survey Report, (since, 1976).
- Ministry of Population and Health. (2011). Nepal Population Report 2011.
- Central Bureau of Statistics. (1995, 2003). Population Monograph of Nepal, Kathmandu (since 1995).
- Central Bureau of Statistics. (2011). Census Reports, (since 1911)
- Central Bureau of Statistics. (1995, 2004, 2011). Nepal Living Standard Survey Reports, Kathmandu Nepal.

Course Title: Statistical Computing-II **Full Marks: 100**

(Practical Paper) **Pass Marks: 50**

Course Number: STA528 **Total Credits: 4**

Total Lectures: 21-23 (3 Hours / Lecture)

Total Duration: 64 hours

Course Objective: The objective of this course is to enable the students to apply theories learnt in solving statistical problems.

SN	Subject Area	No. of Practicals
1	Mathematical Demography	5
2	Sampling Theory	5
3	Design of Experiments	5
4	Two optional Papers of Semester-II	10
	Total	25

Semester III

Course Title: Bayesian Inference **Full Marks: 75**

Course Code: STA631 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objectives: The objective of this course is to make the students familiar with the Bayesian paradigm and to impart knowledge of Bayesian methods for inference, computational methods for posterior summaries and Bayesian regression models.

UNIT 1: Elements of Bayesian Paradigm 8 hrs

- Quantification of uncertainty; Bayesian definition of probability, subjectivity and objectivity; Bayes' theorem, Prior, Likelihood and Posterior
- Prior and prior distributions:
  - Non-informative priors, Improper priors, Conjugate priors and Elicited priors
  - Some special types of priors: Jeffrey's prior, Hartigan's prior and Maximum entropy prior

UNIT 2: Bayesian Inference 18 hrs

- Fundamentals of Bayesian Inference
  - Concepts of Likelihood, Kernel, sufficiency, MLE, Independence and exchangeability
  - Methods of combining prior information with data; Posterior and predictive distributions
  - Conjugate analysis of the cases of binomial, Poisson and normal samples
- Bayesian estimation including posterior conditioning, credible region and HPD, Bayesian inference for normal distribution, predictive distribution
- Hypothesis testing, posterior odds ratio and Bayes' factor
- Bayesian decision rule, Utility and loss functions, Bayes risk, Point Bayes estimates under various loss functions, Lindley's paradox

UNIT 3: Computational Methods in Bayesian inference 8 hrs

- Simulation-based computation: Methods of generating independent sample from distribution; IID sampling; Rejection sampling
- Basic Monte Carlo integration and Importance sampling
- Introduction to Markov chain Monte Carlo (MCMC) methods: Metropolis-Hasting algorithm, Gibbs Sampling and their user-friendly implementation.

UNIT 4: Regression Models from the Bayesian Perspective 7 hrs

- Development of the linear regression model, posterior distribution for the model parameters Conjugate prior analysis for bivariate regression models, random intercept model and random coefficient model
- Bayesian Hierarchical ansd Mixture Models models: Formulation, selection and diagnostics;

UNIT 5: Model specification and checking 4 hrs

- Model selection as a decision problem, Bayesian cross-validation as an approach to diagnostics
- Deviance information criterion (DIC) and Bayesian information criteria (BIC).

Review of the overall course 3 hrs

Reference Books:

1. Box, G.E.P. Bayesian inference in Statistical Analysis
2. Phillips, L. D. Bayesian Statistics for social sciences, Brunai university
3. Carlin, J. B. and Louis T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*, second edition. New York: Chapman & Hall.
4. Congdon, P. (2001). *Bayesian Statistical Modelling*. Chichester: John Wiley & Sons.
5. Spiegelhalter, D. J, Thomas, A., Best, N., and Lunn, D. (2003). *WinBUGS Version 1.4 User Manual*. UK: MRC Biostatistics Unit, Cambridge,.
6. Lindley, D.V. (1980). *Making Decisions*, 2nd ed. New York: Wiley
7. O'Hagan, T. (1998). Bayesian Inference (Vol. 2B of *Kendall's Advanced Theory of Statistics*), UKn: Arnold
8. Lee, P.M. (2004). *Bayesian Statistics: An Introduction*, 3rd edition, London: Arnold.
9. Bernardo, J.M. and Smith, A.F.M. (1994). *Bayesian Theory*. New York: Wiley

Bayesian **Course Title: Research Methodology** **Full Marks: 75**

Course Code: STA632 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: The objective of this course is to make students familiar with research techniques in social sciences. After completion of this course, the students will be able to carry out research work independently.

Unit 1: Introduction to Social Research Methodology 20 hrs

- Scientific Approach of Research
    - Concept and nature of social researches
    - Process of scientific enquiry, planning of social researches and formulation of hypotheses/research problems
    - Setting goals of research problems
    - Criterion of good research problems and statement of problems
    - Principle step of solving research problems
    - Significance of research problems and hypotheses
    - Generality and specificity of problems and hypotheses
    - Multivariate nature of behavioral research
- Research Design
    - Concept and meaning of research design
    - Dimension of research design
    - Purposes and principles of research design
    - Function of research design
    - Research design process
    - Criteria of good research design and inadequate research designs
- Types of Researches
    - Types of researches
    - Social scientific research
    - Ex-post-facto research
    - Laboratory experimental research
    - Field experimental research
    - Field studies research
    - Survey research
    - Case study research
    - Action research
    - Participatory action research

- Qualitative and Quantitative Researches
    - Origin of qualitative and quantitative researches
    - Collection of qualitative information and their analysis
    - Measurement of quantitative/qualitative variables
    - Types of qualitative research
- Methods of Data Collection
    - Observation, interview, questionnaire and schedules
    - Nominal group technique
    - Delphi method
    - Focus group discussion
    - Snowball sampling method

Unit 2: Measurements and Scales 12 hrs

- Measurements and Scales
    - Concept of measurement and scale
    - Nominal, ordinal, interval and ratio measurement scales
    - Standard score, σ , T and Percentile scores
- Reliability and Validity
    - Concept of reliability and validity
    - Test of reliability
    - Content validity, criterion related validity and construct validity
    - Measure of validity
    - Estimation of true score of the test
- Social scales
    - Scales used in measuring-mental health, stress and strain, and life changes experiences
    - Social support, social conflict and work-family conflict
    - Epidemiological depression and social reaction
    - Concept of sociometry
    - Semantic differential and Q-method

Unit 3: Sample Designs, Plans, Data Analysis and Report Presentation 13 hrs

- Sample Designs and Plans
    - Sample plans and designs
    - Selection of optimum size of sample
- Concepts and Techniques of Data Analysis
    - Causal analysis

- Cause and effects analysis
- Canonical analysis
- Factor analysis
- Survival analysis
- Data analysis by using multiple regression analysis
- Multicollinearity and correlation matrix
- Binary logistic regression
- Non-linear regression
- Two-stage least squares
- Multinomial logistic regression
- Probit and logit analysis
- Interpretations and presentation of the results with examples

- Report, Thesis and Research Paper Writing
  - Report writing
  - Thesis writing
  - Research paper writing and research activities
  - Typing of research documents
  - Writing a grant proposal
  - Criteria for a good grant proposal
  - Common shortcomings of grants proposal
  - Some formats and examples of thesis writing, report writing and research paper writing

Review of the overall course 3 hrs

References Books:

1. **Aryal, T.R.** (2008): *Research Methodology*, Paluwa Prakashan Ltd., Kathmandu
2. **Abbas, T. and Charles, T.** (2002): *Handbook of Mixed Methods in Social and Behavioral Research*, Sage Publications
3. **Donna, M. and Pauline, E.G.** (2008): *The Handbook of Social Research Ethics*, Sage Publications
4. **Drapper, N. and Smith, H.** (1968): *Applied Regression Analysis*, John Wiley and Sons
5. **John, F.** (2008): *Applied Regression Analysis and Generalized Linear Models*, Sage Publication Inc
6. **Richardson, J.** (2002): *Handbook of Qualitative Research Methods for Psychology and the Social Sciences*, Blackwell Publishing Co
7. **Kerlinger, F.N.** (1983): *Foundations of Behavioural Research*, Surjeet Publication, India
8. **Kish, L.** (1965): *Survey Sampling*, John Wiley and Sons
9. **Moser, C and Kaltan, G.** (1979): *Survey Methods in Social Investigations*, Heinman Education Books, UK

10. **Pranee, L.R. and Douglas, E.** (1999): *Qualitative Research Methods: A Health Focus*, Oxford University Press
11. **Singh, M.L.** (1999): *Understanding Research Methodology*, Kathmandu.
12. **Strauss, A. and Corbin, C.** (1998): *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, Sage Publication

Course Title: Biostatistics **Full Marks: 75**

Course Code: STA633 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: The objective of this course is to impart knowledge of epidemiology, clinical trials and application of biostatistical techniques in handling health related data.

UNIT 1: Epidemiology

20 hrs

Basic epidemiologic concepts and principles, concept of rates, ratios, incidence and prevalence
Study designs: Types of study design in clinical research - cross-sectional, case-control, cohort studies, experimental studies, ecological studies, choice of study designs

Analysis of epidemiological studies: Issues in analysis of epidemiological data- selection, bias, confounding and interaction, application of multiple logistic regression in epidemiological data
Epidemiology in disease control- screening tests

UNIT 2: Clinical Trials

10 hrs

Introduction: Experimental study design and its importance, randomization and blinding
Ethical issues in clinical trials
Conduct of clinical trials - single centric and multi-centre trials
CONSORT guidelines, role of data safety and monitoring board (DSMB) in conducting clinical trials

UNIT 3: Survival Analysis

15 hrs

Introduction: History and development of survival analysis, need and importance of survival analysis over the standard statistical analysis techniques, concept of event, censoring, right censoring, left censoring, interval censoring, reasons of censoring, structure of time to event data.

Estimation of Survival Functions: Survival function, hazard function, cumulative hazard function, Kaplan-Meier(K-M) estimate of survival function, life table estimate of survival function, Nelson-Aalen estimate of the survival function, Kaplan -Meier estimate of the hazard function, estimation of median and percentiles of survival times, construction of K-M survival curves and interpretations

Comparison of survival experiences: Comparison of survival experiences between two or more groups of survival data - Log-rank test, Gehan's generalized Wilcoxon test, Tarone-Ware test, Peto test

Regression model in survival analysis: Need of regression analysis in handling time to event data, Cox Proportional Hazards (PH) Model with one and several covariates.

Review of the overall course 3 hrs

Reference Books:

1. Hennekens, C.H. and Buring J.E. (1987): Epidemiology in Medicine, Edited by Sherry L. Mayrent. Little, Brown and Company Boston, Massachusetts 02108.
2. Rothman, K. J. and Greenland, S (1998): Modern Epidemiology, Lippincott Williams and Wilkins.
3. Friedman, L.M., Furburg, C. and Demets, D.L.(1998): Fundamentals of Clinical Trials, Springer Verlag.
4. Mathews, J.N.S.(2006): Introduction to Randomized Controlled Clinical Trials, Chapman and Hall/CRC, New York
5. Collett, D. (2003): Modeling Survival Data in Medical Research, Chapman and Hall/CRC, New York
6. Hosmer D.W. and Lemshow, S. (1999): Applied Survival Analysis: Regression Modeling of Time to Event Data, John Wiley and Sons, New York.

Course Title: Environmetrics **Full Marks: 75**

Course Code: STA634 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: The main objective of the course is to acquaint students with statistical methods including advance statistical models incorporating nonlinear models, generalized linear models and some specific environmental models widely applied in environmental and environmental health studies**.**

UNIT 1: Introduction, Environmental and Environmental Health Variables and Studies 10 hrs

Environmetrics: Introduction, origin and its historical development, concept of environmental epidemiology

Environmental Pollution: Definitions and types of environmental pollution and their impacts on human health, climate change and health risks

Environmental and Environmental Health Variables/Statistics: Risk assessment (pollutant exposures, meteorological conditions, etc) variables and outcome assessment (examinations, diagnostic tests, etc) variables and statistics, qualitative and quantitative measures, direct and indirect measures, indicators, sources of environmental and related data, biomarkers, national and international standards of pollutant levels.

Environmental Health Study Designs: Descriptive studies, analytical studies, ecological studies, cohort studies, cross-sectional and longitudinal studies, experimental or intervention studies, case-crossover design, meta-analysis

UNIT 2: Transformations and Generalized Least Squares 8 hrs

Transformations in Models: Variance stabilizing transformations, transformations to linearize models, Box-Cox transformation

Generalized Least Squares: Definition and derivation of generalized least squares, properties

Weighted Least Squares: Definition, Condition under which generalized least square is a weighted least square, choice of weights

UNIT 3: Nonlinear Models 7 hrs

Introduction: Definition, nonlinear functions, differences between linear and nonlinear models, assumptions, their uses in environmental and related studies

Inference: Model specification, nonlinear least squares, linearization (Gauss-Newton) method of parameter estimation, inference in nonlinear models, Pseudo R^2

UNIT 4: Generalized Linear Models 14 hrs

Exponential Family of Distributions: probability density function, moments (mean and variance) and its members (normal, binomial, Poisson, negative binomial, exponential, gamma)

Generalized Linear Models (GLM):

Introduction: Definition, transformations versus GLM, canonical link functions (identity, log, logit, reciprocal links), models for different canonical links including Poisson and logit models and their uses in environment related studies

Inference: Maximum likelihood estimation, iterative re-weighted least square (IRLS) estimation, interpretation of parameter coefficients, tests of significance, model deviance.

Residual Analysis: Types of residuals, raw, deviance and Pearson residuals and statistics, over-dispersion, analysis of deviance, Omnibus test, pseudo R^2

UNIT 5: Environmental Pollution Assessment Models 6 hrs

Air dispersion models: Introduction, Gaussian plume model: model specification, characteristic features, meteorological conditions, dispersion coefficient, plume rise

Dose-response models: Introduction, dose-response curve, functional forms, Hill function model: model specification, estimation of parameters, uses

Review of the overall course 3 hrs

Reference Books:

1. Merrill, R. M. (2010): Environmental Epidemiology, Principles and Methods, Jones and Bartlett India Pvt. Ltd., New Delhi.
2. McCullagh, P. and Nelder, J. A. (1989): Generalized Linear Models, Chapman and Hall, London.
3. Cameroon, A. C. and Trivedi, P. K. (1998): Regression Analysis of Count Data, Cambridge University Press, UK.
4. Montgomery, D. C., Peck, E. A. and Vining, G. G. (2003): Introduction to Linear Regression Analysis, John Wiley and Sons, INC, Singapore.
5. Shrestha, S. L. (2010): Statistical Methods for Environment, Biological and Health Sciences, Ekta Books, Kathmandu, Nepal.

Course Title: Time Series Analysis **Full Marks: 75**

Course Code: STA635 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: The objective of this course is to impart the knowledge of time series analysis and its applications.

UNIT 1: Stationarity 5 Lhr

Stationary and nonstationary time series, tests for stationarity, correlogram, Ljung-Box statistic, unit root test, random walk, trend stationary and difference stationary time series.

UNIT 2: Time series models 10 Lhr

Autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), autoregressive integrated moving average processes (ARIMA), Box_Jenkins (BJ) methodology, autocorrelation and partial autocorrelation functions, estimation of the ARIMA model, diagnostic checking and forecasting.

UNIT 3: Vector Autoregression (VAR) 10 Lhr

Introduction, model specification, assumptions, estimation of VAR, forecasting with VAR, some problems with VAR, application of VAR.

UNIT 4: Exponential smoothing 10 Lhr

Simple and weighted moving averages, exponential moving average, single, double and triple exponential smoothing, smoothing equation, smoothing constant, procedures for estimation.

UNIT 5: ARCH and GARCH models 10 Lhr

Autoregressive conditional heteroscedastic (ARCH) model, ARCH (q) model specification and estimation, lagranges multiplier test, Generalized autoregressive conditional heteroscedastic (GARCH) model, GARCH (p, q) model specification and estimation.

Review of the overall course 3 hrs

Reference Books:

1. **Gujarati, D. N. (1995):** *Basic Econometrics*, McGraw-Hill, Inc.

2. **Walter Enders (2004):** *Applied Econometric Time Series,* John Wiley and Sons.

3. **Ramanathan, B.** (2002): *Introductory Econometrics with Applications,* South-Western Thomson Learning, Singapore
4. **Maddala, G. S.** (2002): *Introduction to Econometrics*, John Wiley and Sons

Course Title: Operation Research **Full Marks: 75**

Course Code: STA636 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Objectives: This course is aimed to enable the students to understand and develop the skill of applying operations research tools and also to impart substantial knowledge of handling model based decision problems.

Unit 1: Linear programming 8 hrs

General Nature of programming problems, Scope and limitation, concepts of feasible, infeasible, optimum solutions, infeasible, optimum solutions, effective, ineffective, simultaneous linear equations, basic solutions, linear transformations, point sets, lines and hyper planes, convex cones.

Unit 2: Formulation of Linear Programming Problems 2 hrs

Simple linear programming problems, elictbalancing problems, blending problems, inder-industry problems.

Unit 3: Graphic Solution of Linear Programming 2 hrs

Maximization, Minimization, bounded, unbounded solution.

Unit 4: Simplex Method 7 hrs

Slack, surplus and artificial variables: Theory of simplex method theory and applications of reduction of any feasible solution to basic feasible solution improving basic feasible solution, unbounded solutions, optimality conditions, degeneracy and breaking ties, inconsistency and redundancy, tableaus format of simplex computations, and its use conversion of minimization into minimization. With examples, solution of simple methods when artificial variables included.

Unit 5: Duality Theory and its Ramifications 5 hrs

Dual linear programming problems, fundamental properties of dual problems, complementary slackness, unbounded solution in the primal, dual-simplex algorithm,

Unit 6: Post optimal / Sensitivity Analysis 5 hrs

Post-optimality problems, changing the price vector, changing the requirement vectors, adding variables or constraints, upper and lower bounds.

Unit 7: Integer Programming 3 hrs

Introduction, Application of integer programming Formulation possibilities through mixed integer programming, Methods of integer programming, Branch and bound algorithm, Gomery Fractional cut algorithm.

Unit 8: Transportation and Assignment Problem 4 hrs

Introduction, north-west, least last, Vogel's approximation method. Solution of transportation problem by Stepping method and MODI method, Duality and degenerate transportation problem.

Unit 9: Inventory Models 5 hrs

Introduction, deterministic models: No shortage, shortage allowed, finite shortage cost but variable demands and inputs (discrete and continuous stocks)

Unit 10: Sequencing Model 4 hrs

Introduction, Problems Assumptions, Processing of *n* jobs through one machine, two machines processing *n j*obs through *m* machines processing two jobs through *m* machines.

Review of the overall course 3 hrs

Reference Books:

1. Bernard W. Taylor III (2009): Introduction to Management Science, Prentice Hall, India.
2. Hadle, G. (1978): Linear Programming, Edision-Wesley Publishing Co.
3. Gupta Prem Kumar, Hira D. S. (2007): Operations Research, 4th edition S. Chand & Company Ltd.
4. Paul T.J. James (1996): Total Quality Management: An Introductory Text, Prentice Hall
5. Sthapit Azaya et al (2010); Data Analysis and Modeling, Asmita Publication, Kathmandu.
6. Vohra N.D. (2006): Quantitative Techniques in Management TATA McGraw Hill.

Course Title: Survival Analysis **Full Marks: 75**

Course Code: STA637 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: The objective of this course is to impart knowledge of different survival modeling techniques such as semi-parametric, parametric and accelerated failure time models with special focus to clinical data.

UNIT 1: Review of Basic Survival Analysis Terms and Techniques 5 hrs

Need and importance of survival analysis in clinical research, concept of event, censoring, reasons of censoring, estimation of Kaplan-Meier (K-M) survival functions, hazard functions, survival times, K-M survival curves, Log-rank test.

UNIT 2: Cox Proportional Hazards Regression Model

15 hrs

Modeling the hazard function, Cox Proportional hazards (PH) model, linear component of Cox PH model, fitting of Cox PH model, confidence intervals and hypothesis tests for regression coefficients, strategy for model selection, interpretation of parameter estimates, estimation of hazards and survival functions, assumptions of Cox PH model, tests of proportionality of hazards assumption- graphical method, test based on Schoenfield residuals, interaction with time.
Assessment of model adequacy of Cox regression model: Residuals for the Cox regression model, assessment of the model fit, identification of influential observations, overall goodness of fit
Extended Cox regression model: Stratified proportional hazards model, time varying covariates.

UNIT 3: Parametric Proportional Hazards Model 10 hrs

Parametric proportional hazards model: Exponential distribution, Weibill distribution; Assessing the suitability of a parametric model, fitting of a parametric model to a single sample, model for the comparison of two groups, Weibull proportional hazards model, fitting of Weibull proportional hazards model, log-linear form of the model, Gompertz proportional hazards model, tests of proportionality of hazards assumption in parametric PH models, residual analysis, goodness of fit of the model, model selection

UNIT 4: Accelerated Failure Time Models

15 hrs

Probability distributions for survival data: Log-logistic distribution, lognormal distribution, gamma distribution, inverse Gaussian distribution
Exploratory analysis for the selection of appropriate model
Accelerated Failure Time(AFT) models: Concept of AFT, AFT model for comparing two groups, general AFT model, log-linear form of AFT model, interpretation of parameter estimates and measures, difference between PH metric and AFT metric in survival analysis
Parametric AFT models: Weibull AFT model, Log-logistic AFT model, lognormal AFT model
Residual analysis for parametric models: Standardized residuals, Cox-Snell residuals, deviance residuals, score residuals.

Review of the overall course 3 hrs

References Books:

1. David Collett (2003): Modeling Survival Data in Medical Research, Chapman and Hall/CRC, New York.
2. John P. Klein & Melvin L. Moeschberger(2003): Survival Analysis Techniques for Censored and Truncated Survival Data, Springer Publication
3. David.W Hosmer and Stanley Lemshow (1999): Applied Survival Analysis: Regression Modeling of Time to Event Data, John Wiley and Sons, New York.
4. Terry M. Therneau and Patricia M. Grambsch(2001): Modeling Survival Data: Extending the Cox Model, Springer Publication.
5. Jerald F. Lawless (2003): Statistical models and Methods for Lifetime Data, John Wiley & Sons Inc Publication.

Course Title: Actuarial Statistics **Full Marks: 75**

Course Code: STA638 **Pass Marks: 37.5**

Total Credits: 3

Total Lecture Hours: 48

Course Objective: To impart knowledge and improve level of understanding of insurance and actuarial statistics along with their applications.

UNIT 1: Insurance **15 hrs**

Introduction of actururial sciences, nature and functions of insurance, benefits and costs of insurance system to the society; economic theories of insurance; the mathematical basis for insurance; insurable interest.

Life Insurance: Essential features of life insurance contract; Risk selection for life insurance; Sources of risk information.

Life annuities: Single payment, continuous life annuities, discrete life annuities, life annuities with monthly payments, commutation functions, varying annuities, recursions, complete annuities-immediate and apportionable annuities-due.

Health insurance: Types of health insurance coverage; Exclusion in health insurance policies. payment of claim.

Other schemes of insurance.

UNIT 2: Actuarial Statistics **30 hrs**

The economics of insurance, utility theory, application of probability to problems of life and death, determination of single premiums for insurances and annuities, theory and practice of pension funding, assumptions, basic actuarial functions and population theory applied to private pensions.

Survival distributions and life tables, life insurance, life annuities, net premium, premium series, multiple life functions, multiple decrement models, valuation theory for pension plans, the expense function and dividends.

Risk and Mortality Table: Mortality tables and its classification; construction of mortality tables; premium calculation of various life policies.

Exposure formulas: Techniques of calculating exposures from individual records including consideration involving selection of studies, various observation periods and various methods of tabulating deaths, techniques of calculating exposures from variation schedules, use of interim schedules and variations in observation period or method of grouping deaths and practical aspects of construction of actuarial tables.

Review of the overall course 3 hrs

Reference Books:

1. **Dorfman, Mark.S (1991)**: *Introduction to Risk Management and Insurance*, Prentice Hall, India
2. **Mishra, M.N. (1989):** *Principles and practice*, S.Chand and Company, India
3. **Atkinson, M.E. and Dickson, D.C.M. (2000).** *An Introduction to Actuarial Studies*, Elgar Publishing.

Course Title: Statistical Computing-III **Full Marks: 100**

(Practical Paper) **Pass Marks: 50**

Course Number: STA639 **Total Credits: 4**

Total Lectures: 21-23 (3 Hours / Lecture)

Total Duration: 64 hours

Course Objective: The objective of this course is to enable the students to apply theories learnt in solving statistical problems.

SN	Subject Area	No. of Practicals
1	Bayesian Inference	5
2	Research Methodology	5
3	Three Optional Papers of Semester-III	15
	Total	25

Semester IV

Course Title: Dissertation **Full Marks: 100**

Course Code: STA641 **Total Credits: 4**

Supervision: 1 hour / week

The guidelines and format of dissertation will be decided by the department.

Course Title: Meta Analysis **Full Marks: 50**

Course Code: STA642 **Pass Marks: 25**

Total Credits: 2

Total Lecture Hours: 32

Course Objective: The course has been designed with the aim of enabling the students to understand the basic principles of, and to apply, the different methods of Meta Analysis.

UNIT 1: Introduction 4 hrs

Introduction to Meta Analysis; Development and uses, Systematic reviews, characteristics of systematic review, individual studies, the summary effect, heterogeneity of effect sizes, the streptokinase meta-analysis, statistical significance, clinical importance of the effect, consistency of effects

UNIT 2: Effect Size and Precision 4 hrs

Treatment effects and effect sizes, parameters and estimates, outline of effect size computations, raw (un-standardized) mean difference(D), standardized mean difference(d and g) response ratios.

Effect sizes based on binary data (2×2) tables: Risk ratio, odds ratio, risk difference, choosing an effect size index; Effect sizes based on correlations, factors affecting precision

UNIT 3: Assessing between Study Heterogeneity 7hrs

Hypothesis tests for presence of heterogeneity: Standard χ^2 test, extensions/alternative tests.

Graphical informal tests/Explorations of heterogeneity: Plot of normalized (z) score, Forest plot, Radial Plot (Galbraith diagram), L'Abbe plot; possible causes of heterogeneity

Methods of investigating and dealing with sources of heterogeneity: Changing the scale of outcome variable, include covariates in regression model, exclude studies, use of random and fixed effect models; Validity of pooling studies with heterogeneous outcomes

UNIT 4: Fixed Effect versus Random Effect Models 7 hrs

True effect size, impact of sampling error, performing a fixed-effect meta-analysis and random effect meta-analysis, definition of a summary effect, estimating the summary effect, extreme effect size in a large study or a small study, confidence interval, model selection

UNIT 5: Publication Bias 6 hrs

Evidence of publication and related bias, seriousness and consequences of publication bias for Meta analysis, predictors of publication bias

Tools to identify publication bias in Meta analysis: The funnel plot, rank correlation test, linear regression test and other methods; 'Rosenthals's file drawer' method, 'Trim and Fill' method

UNIT 6: Reporting the Results of Meta Analysis 2 hrs

Overview and structure of a report, graphical displays for reporting the findings of a Meta analysis.

Review of the overall course 2 hrs

References Books:

1. Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein H. R. (2009). Introduction to Meta-Analysis. West Sussex, UK: Wiley.
2. Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). The Handbook of Research Synthesis and Meta-Analysis (2nd Edition.). New York, NY: Russell Sage Foundation.
3. Sutton, A. J., Abrams, K. R., Jones, D. R., Sheldon, T. A., & Song F. (2000). Methods for Meta-Analysis in Medical Research. John Wiley & Sons, Ltd.
4. Lipsey, M. W. & Wilson, D. (2000). Practical Meta-Analysis. Sage Publications.

Course Title: Nonparametric and Categorical Data Modeling **Full Marks: 50**

Course Code: STA643 **Pass Marks: 25**

Total Credits: 2

Total Lecture Hours: 32

Course Objective: To impart knowledge, understanding and uses of nonparametric regression models and statistical models for categorical response variables

UNIT 1: Nonparametric Regression Models 15 hrs

- Parametric versus nonparametric regression, smoother, smoothing parameter, scatterplot smoother
- Bin smoothers
- Local averaging: nearest neighborhood, running mean and running line smoothers
- Kernel estimation: Locally weighted averaging, Kernel functions, weights and smoothing
- Locally weighted regression smoother (Loess): *k* nearest neighborhood, span, tri-cubic weight function
- Regression splines: piecewise polynomials, interior knots, smoothing function

Illustrative Examples

UNIT 2: **Regression Models for Categorical Responses** 15 hrs

- Binary Logistic Regression Model (Review only)
- Multinomial Logistic Regression Model
  Model specification, assumptions, estimation of parameters (derivation not required) with interpretations, examples of fitted models with model adequacy tests
- Ordinal Logistic Regression Model
  Model specification, assumptions, estimation of parameters (derivation not required) with interpretations, examples of fitted models with model adequacy tests

Review of the overall course 2 hrs

Reference Books:

1. Hastie, T. J. &Tibshirani, R. J. (1990) Generalized Additive Models, Chapman and Hall /CRC, USA.
2.Agresti, A. (1990). Categorical Data Analysis. New York: Wiley and Sons, Inc.
3. Greene, W. H. (2003): Econometric Analysis (fifth edition), Pearson Education Inc., Singapore.
4. Montgomery, D. C., Peck, E. A. and Vining, G. G. (2003): Introduction to Linear Regression Analysis, John Wiley and Sons, INC, Singapore.

7. Evaluation Scheme

The student performance will be basically judged through attendance and examination. Different modes of evaluation system are given as follows.

- Written examinations
- Oral (Viva-Voce) examinations
- Presentations for theoretical papers
- Submission and Presentation of assignment work
- Thesis, presentation and Viva-Voce for thesis work

The guidelines for evaluation are as follows.

- A minimum of 80% attendance will be required for students to appear in final examination.
- Internal assessment covers 40% of the total marks for each theory papers, computational papers and dissertation.
- Internal assessment marks includes marks of two written exams (mid-term, pre-board), class performance and attendance and at least two of the following: assignment work, class seminar, presentation, oral examination, class test in each of the papers.
- A final examination will be conducted for each of the papers as per the total marks and marks secured by the students will be converted to 60% of the total marks.
- An initial presentation will be required for the proposed dissertation title.
- A pre-submission seminar will be required for the submission of the dissertation.
- Dissertation will be evaluated through internal assessment and on the basis of external expert examination followed by Viva-Voce.

A breakdown of marks for assessment for each of the courses and dissertation work is given below.

Theoretical Papers

Nature of Evaluation	Examination	Allocated Marks (for 3 Credits)	Allocated Marks (for 2 Credits)
Attendance and Class Performance	Evaluation	5	3
Assessment	Mid-term (written)	5	3
	Pre-board (written)	10	7
	Assignment work / Oral test / class test Presentation / class seminar	10	7
Total of Internal Assessment		30 (40% of Total)	20 (40% of Total)
Final Examination	Written	45 (60% of Total)	30 (60% of Total)
Total		75	50

Computing Papers (4 Credits)

Nature of Evaluation	Examination	Allocated Marks
Internal		
Practical class Performance + Attendance	Evaluation	10
Submission of practical assignments	Evaluation	30
Total of Internal Evaluation		**40**
External		
Final Practical Examination	Written	40
	Viva-Voce	20
Total of Final Examination		**60**
Total of the Practical Evaluation		**100**

Dissertation (for 4 Credits)

Nature of Evaluation	Examination	Allocated Marks
Internal		
Regularity and Research Work	Evaluation	10
Perseverance	Evaluation	10
Pre-submission Assessment (Presentation, Compilation, Documentation)	Evaluation	20
Total of Internal Evaluation		**40**
External		
Dissertation Evaluation	Evaluation	30
Viva-Voce	Evaluation	30
Total of External Evaluation		**60**
Dissertation Evaluation Total		**100**