SYLLABUS
FOR POST GRADUATE DIPLOMA IN STATISTICAL TECHNIQUES AND COMPUTATION (PGDSTC)COURSE

NODAL DEPARTMENT : DEPARTMENT OF STATISTICS

DIBRUGARH UNIVERSITY

Syllabus recommended

By

The Board of Studies (BOS) in Statistics, Dibrugarh University in its Meeting held on 02/04/2007 and approved by the Post Graduate Board in its Meeting held on 15 / 05 /2007 and passed by the Academic Council in its Meeting held on 22 / 05 / 2007

DIBRUGARH UNIVERSITY

ASSAM

STRUCTURE OF THE COURSE

Post Graduate Diploma in Statistical Techniques and Computation (PGDSTC) course

Paper / Paper Code	Title	Contact hours per week	Distribution of Marks		
			End Term	Internal Assessment	Total
Semester – I					
Paper – 1.1	Basic Statistics – I	4 Lectures	60	15	75
Paper – 1.2	Applied Statistics – I	4 Lectures	60	15	75
Paper – 1.3	Statistical Computing – I	4 Lectures / Demonstration	60	15	75
Paper – 1.4	Statistical Computing – II	4 Lectures/ Demonstration	60	15	75
Semester – II					
Paper – 2.1	Basic Statistics – II	4 Lectures	60	15	75
Paper – 2.2	Applied Statistics – II	4 Lectures	60	15	75
Paper – 2.3	Statistical Computing – III	4 Lectures/ Demonstration	60	15	75
Paper – 2.4	Project : Report writing + Presentation + Viva-Voce Project shall be based on primary data/ secondary data / and or both * Details (refer to Annexure – III in Regulations for PGDSTC)				60 + 15 = 75

SYLLABUS FOR P.G.D.S.T.C. COURSE

Paper 1.1 : Basic Statistics - I   **Marks – 60 (T.E.) + 15 (I. A.) = 75**

Prerequisites : Elementary Mathematical Tools.
Statistics : Definition and limitation; statistical data; types of data: primary & secondary; sources of data; methods of data collection; preparation of questionnaire & schedule; Tabulation of data. Diagrammatic representation of data; classification of data & construction of frequency distributions; graphical representation of data.

Measures of location, dispersion & relative dispersion. coefficient of variation; moments, measures of skewness & kurtosis.

Bivariate data; scatter diagram; bivariate frequency distribution;

The correlation coefficient for a bivariate frequency distribution; Product moment correlation coefficient and its interpretation. Spurious correlation and its interpretation.

Two variable linear regression model, description of the data for two variable linear regression model, estimating regression coefficients. ANOVA in regression, interpretation of regression parameter and other measures.
Quantitative vs qualitative data;
Various scales of measurements : nominal, ordinal, interval and ratio scale.
Attribute : Measures of association and contingency.

References
   1. Statistical Methods- an introductory Text –Medhi, J : New Age International.
   2. Basic Statistics: Goan, Gupta and Dasgupta, World Press
   3. Ranga Swamy : A Text Book of Agricultural Statistics, New Age
   4. Draper & Smith : Applied Regression analysis, John Wiley.
   5. Chaterjee & Price : Regression analysis by example, John Wiley .

Paper 1.2 : Applied Statistics - I   **Marks – 60 (T.E.) + 15 (I. A.) = 75**

Sample Survey : Concepts of population, sample, parameter, statistic, census, sample survey, Sample Survey Vs Complete enumeration, Planning and execution of sample survey (various stages of sample survey), errors in survey, purposive and judgment sampling :

Techniques of Sample Survey : Simple Random Sampling ;Stratified Random Sampling; Systematic Sampling ; Multistage Sampling; Cluster Sampling.

Design of questionnaires and schedules for data collection to conduct surveys in different fields of studies.   (Marks – 25)

Analysis of Variance and Design of Experiments :- AOV of one way and two way classified data : Basic Principles of Design of Experiments; Completely Randomized Design (CRD); Randomized Block Design (RBD); Analysis of Design with missing observation.

Design of experiment for laboratory and field experiments in agriculture, industries, bio-medical sciences etc.   (Marks – 25)

Demography : Sources of demographic data; some important national level surveys : SRS, NFHS, CSO; NSSO); Indian population census. Various measures of fertility, mortality life table : basic idea and application in actuarial sciences; growth model : Logistic, Makeham's method of population projection.

Design of questionnaire and schedule for collection of demographic data.   (Marks – 25)

References :

1. Cochran, W. G.(1977) : Sampling Techniques 3rd Ed., John Wiley & Sons.
2. Deming, W.E. [1960] : "Some Theory of Sampling" John Wiley & Sons.
3. DesRaj & Chandhok , P [1990] : Sample Survey Theory, Narosa Publishing House.
4. Medhi, J. [1998] : Statistical Methods : An introductory Text New Age International (P) Ltd.
5. Singh D and Choudhuri F.S. [1995] : Theory and Analysis of Sample Survey, New Age International (P) Ltd.
6. Sampath S. [2001] : Sampling Theory and Methods, Narosa Publishing House.
7. Kemthorne, O : Design and Analysis of Experiments
8. Cochran & Cox : Design of Experiments
9. Ram Kumar, R : Technical Demography, New Age
10. K. Srinivason : Basic Demographic Technique and Application, SAGE publication

<u>Paper 1.3 : Statistical Computing - I</u>   **Marks – 60 (T.E.) + 15 (I. A.) = 75**
Prerequisites : Exposure to computer.
Database construction, preparation of data sheets, tabulation, graphical presentation, descriptive statistical summarization, analysis and interpretation of data using the techniques introduced in paper 1.1, exploiting MS-Excel, Statistical and other software packages.

<u>Paper 1.4 : Statistical Computing - II</u>   **Marks – 60 (T.E.) + 15 (I. A.) = 75**
Analysis and interpretation of data using the techniques introduced in paper 1.2, exploiting MS-Excel, Statistical and other software packages.

<u>Paper 2.1 : Basic Statistics - II</u>   **Marks – 60 (T.E.) + 15 (I. A.) = 75**
The role of the hypothesis in scientific investigations. Introduction of general theory of testing of Statistical hypotheses, p values.
Parametric Testing Procedures : Large sample tests for proportions and means. χ^2 - test for : goodness of fit, independents of attributes single population variance, homogeneity of population variances. t-test for : single mean, difference of two means, difference of means of paired observations, regression coefficient, F-test for : equality of population variances, multiple correlation coefficient, linearity of regression, equality of several means.   (Marks – 30)
Non-parametric Procedures : Sign test, signed - rank test; Wilcoxon test, Mann- Whitney test, t-test for the significance of the difference means; Welch-test for the significance of the difference between mean when the scale parameters and populations are unknown; Kolmogorov – Smirnov test for the homogeneity of two or more frequency distributions   (Marks – 30)
Multiple Regression Analysis : Introduction to multiple regression, relevant formulae, Interpretation of parameters and of various measures associated with it.   (Marks – 15)

 References :

1. Hogg. & Craig : Introduction to Mathematical Statistics, Pearson Education
2. Siegal. S : Non Parametric Statistics for behavioural Sciences, McGraw Hill
3. Gibbon J.D. : Non- parametric Statistical Inference, McGraw Hill
4. Draper & Smith : Applied Regression analysis, John Wiley.
5. Chaterjee & Price : Regression analysis by example, John Wiley .

Paper 2.2 : Applied Statistics - II   **Marks – 60 (T.E.) + 15 (I. A.) = 75**

Time series analysis and forecasting : The component of a time series. Forecasting using smoothing methods. Forecasting time series using trend projection. Forecasting time series using classical decomposition. Forecasting time series using regression models.   (Marks -25)

Industrial Quality Control (IQC) :Statistical Process Control (SPC), Seven tools of Quality improvement, Basic concepts of SQC, Basic SQC tools (Shewhart's Control Charts) for variables and attributes. Acceptance Sampling Plans for variables / Attributes (single sampling only). Quality cycles (Plan – Do – check – Act cycle)   (Marks -25)

Multivariate Analysis : Introduction to Multivariate Analysis and Application of Multivariate Techniques in Data Analysis, Analysis and Interpretation of results .   (Marks -25)

References :

1. Johnson, R. A. and Wichern D.W :- Applied Multivariate Statistical analysis, Pearson Education.
2. Anderson T.W : An introduction to Multivariate Statistical analysis, Wiley
3. S.K. Ekambarm : Acceptance Sampling
4. Montgomery, D.C. [2009] : Introduction to Statistical Quality Control : A modern Introduction (Sixth Edition), John Wiley & Sons.
5. Eugene L. Grant, Richard S. Leavenworth : Statistical Quality Control, McGraw Hill (1996) Seventh Edition.
6. Bx, Jenkins, Reinsel : Time Series Analysis forecasting and control, prentice Hall.
7. Granger C.W.J : Forecasting in Business and Economics, Academic Press.

Paper 2.3 : Statistical Computing – III   **Marks – 60 (T.E.) + 15 (I. A.) = 75**

Analysis of data and interpretation of the output from analysis using the techniques introduced in paper 2.1 and 2.2, exploiting MS-Excel, Statistical and other software packages

<u>Paper 2.4 : Project Work</u>   **Marks – 60 (Report) + 15 (Viva-Voce) = 75**

Guidelines for the Project Work : PGDSTC course.

1. The project work shall be spread over the entire 2nd semester
2. The project work shall be based on primary / secondary data
3. The assessment will be based on the presentation of the project report plus viva-voce test. Viva-Voce Test is open .

A faculty member assigned by the DMC of the Department shall supervise the project prepared by the student. There shall be an external examiner and an internal examiner (the supervisor) for the evaluation of the project work.

The project work should be chosen such that there is enough scope to apply and demonstrate the statistical techniques learnt in the course.

A project report shall clearly state the problem addressed, the methodology adopted, the assumptions and hypotheses formulated, references consulted, statistical analyses performed and the conclusion drawn.

********* X **********

Rationale

The subject Statistics as a method of decision making in the face of uncertainty on the basis of facts and figures is one of the growing subjects world over and is proved to be the most sought for subject after the breakthrough of Information Technology. The department of Statistics, D.U. continuing its saga of statistics education is pleased to introduce the PGDSTC with a primary view to producing graduates to take fullest possible advantages of statistical methodology in solving statistical problems in industries, in research problems of different fields and of the computing environment that they are likely encounter in their careers. This goal is tried to be achieved through the plan as follows.

We wish to ensure a firm understanding of the basic, and technical elements of statistical computing. The proposed one shall instill a strong knowledge of broad spectrum of the subject ranging from numerical analysis to interactive computer graphics, data banking to data analysis and stress on calculative accuracy and efficiency.

We wish to encourage creativity in solution process by developing a level of confidence that comes only from a full confidence of computing intricacies and exposure of to-day's software offerings. Thus the graduate shall be able to obtain multiple problem solution ranging from " quick and dirty" approximation to elaborate, computationally intense " works of art".

We shall provide a broad perspective and show students how to take a data-based problem from "start to finish". This includes the ability to respond to :

* How do I capture and transport the needed data ?

* What "software and computation" combination is required to solve the problem routinely in future ?

* How can statistical methods and thinking be transferred to other fields ?

* How can students be made "jack of all trades" plus "a master of few" ?

Recognizing that to-day's graduates will be spending most of their professional careers in twenty first century, we shall try to provide not only an understanding of current computing environment but also a generic foundation that will allow graduates to quickly grasp and address tomorrow's needs.

We wish to ensure the Diploma curriculum in statistical computing for those engaged in industry and research and for those who plan to enter an industry or plan to join research. Main emphasis would be given on survey design in bio-sciences, in various fields of social, physical & life science, business, industry health care etc. Questionnaire construction, interview techniques, quantification and measurements of information would be also taught in the course.

We shall provide training to formulate the problems to develop a simplified series of operational steps that lead to a good solution, to teach them how to use statistics to solve problems rather than teaching them to be statisticians.

We shall provide the clients the training to be nimble problem solver with first –rate statistical skills. Solid thinking of statistics and computing is required for integrated solution to many ways of toughest problems in Industries, Finance and Management Institutions and Planning and execution of projects in Govt. agencies and NGO's.

We shall try to reduce the gap that prevails between the statisticians and the personnel engaged in Industry, Research Projects Govt. agencies, NGO's , Planning and Execution, Surveys – Data collection, Data Analysis in Manufacturing, Agriculture, Bio-medical sector etc.

The way subject is formulated, and will be taught, will be made palatable to customers – Engineers, Scientists and Researchers etc. It would be made user friendly. The focus / attention would be given on the needs of the customers.

What for the Course is Designed

To promote statistics and to transfer statistical methods and thinking to other disciplines.

To focus the diverse applications of statistics in real world phenomena and to offer students different perspectives about use of statistics and to focus how statistical methods and designs can be made palatable to personnel in social life & physical sciences, engineering, manufacturing, management and financial institutions and researchers engaged in different projects, Govt. organization projects and that of NGO's.

To focus how statistical tools are important in decision making today and how we can do better even more in future.

To focus how important the statistical thinking and statistical computation are and why it is required for integrated solution to many toughest problems in the real world activities [viz. Industry, Financial and Management Institutions, Planning and Execution of projects in Govt. agencies and NGO's].

To focus on how to transform data into information, information gathering to information processing.

To focus / stress on consulting process by making both clients and statisticians active.

Client	Active	Statistician as a helper	Statistician as a colleague
	Passive		Statistician as a leader
		Passive	Active

Role of clients and statisticians

Why Statistical Computation

```
┌─────────────────┐
│   Problem as    │
│     Posed       │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│      Real       │
│    Problem      │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│   Statistical   │
│   Formulation   │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│  Mathematical   │
│    Solution     │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│  Interpretation │
│  of solution in │
│     words       │
│  understandable │
│   to clients    │
└─────────────────┘
```