Rs - 10 - -

॥ अंतरी पेटवू ज्ञानज्योत ॥ - 2

North Maharashtra University, Jalgaon

Syllabus for S.Y.B.Sc.

STATISTICS

W.E.From June, 2003

॥ अंतरी पेटवू ज्ञानज्योत ॥

NORTH MAHARASHTRA UNIVERSITY, JALGAON

Syllabus for S.Y.B.Sc.

STATISTICS

[W.E.From June, 2003]

Syllabi in Statistics at S.Y.B.Sc. consist of three papers, each paper carries 100 marks as follows:

- Paper I : Continuous Probability Distributions.
- PaperII : Applied Statistics.
- Paper III : Practicals in Statistics.

The syllabi will be effective from June 2003.

The following notations/conventions are used while preparing the syllabi.

μ : Population mean

σ : Population standard deviation

ρ : Population standard deviation

Π : Population proportion

$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu)^2$: *Population mean square*

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$: *Sample mean square*

$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$: *Sample variance*

Paper I: Continuous Probability Distributions

Prerequisites : The knowledge of the topics taught at F.Y.B.Sc. in Statistics.

Objectives : The study of discrete random variable and its probability distribution and related topics was introduced at F.Y.B.Sc. This paper is introduced to continue the study of probability distribution and related topics for continuous random variables. The following objectives are expected for this course:

- Introduction of the concept of continuous random variable and its related topics such as probability density function, distribution function, moment generating function, cumulant generating function, expectation of one and two dimensional random variables, marginal and conditional probability distributions.
- Study of continuous type standard probability distributions, their properties and application in real life situations.
- Study of sampling distributions of some standard statistic.
- Study of tests of significance and their applications in real life situations.

Unit 1. Mathematical concepts [8L,8M]

1.1 Definition of gamma and beta function.

1.2 Different forms of gamma and beta functions.

1.3 Relation between gamma and beta functions (Statement only).

1.4 Expression and discussion of convergence of the following:- e^x, e^{-x}, log(1+x), log(1-x), $(1-x)^{-n}$, $(1+x)^{-n}$, $\sum_{t=0}^{\infty} a^t x^t$ (Statement only).

1.5 Cauchy-schwartz inequality for real numbers (with proof).

1.6 Concept of orthogonal tranformations and Jacobian of transformation.

1.7 Some elementary examples and problems.

Unit 2. Univariate continuous distributions [8L,8M]

2.1 Definition of continuous sample space, continuous random variable, functions of continuous random variable, probability density function, distribution function, statement of properties of distribution function.

2.2 Expection of a random variable, expection of a function of a random variable, moments, raw and central moments, evaluation of moment generating function, cumulant generating function.

2.3 Mode, median and quartiles.

2.4 Transformation of variables, statement of theorem, probability density function of simple monotone functions and probability density function of $Y=X^2$ only.

2.5 Examples and problems.

Unit 3. Bivariate continuous probability distribution [8L,8M]

3.1 Joint probability density function, evaluation of probabilities of region bounded by straight line and circle.

3.2 Marginal and conditional distribution.

3.3 Expection of g(X,Y), moments of bivariate distribution conditional expectation, regression as a conditional expectation, correlation coefficient.

3.4 Probability distribution of function of bivariate random variables using Jacobian of transformation.

3.5 Independence of two random variables, statements of extension to n (>2) random variables.

3.6 Theorems of expectations:- E(X+Y) = E(X)+E(Y) and E(XY)=E(X)E(Y), X and Y independent. Statement of extensions of above theorems to n variables.

3.7 Examples and problems.

Unit 4. Uniform (Rectangular) Distribution. [4L.4M]

4.1 Probability density function

$$f(x) = \frac{1}{b-a} \qquad a < x < b$$

$$= 0 \qquad \textit{otherwise}$$

Notation $X \sim U(a,b)$.

4.2 Distribution function, mean and variance, moment generating function, r^{th} moment about origin.

4.3 Standard form of uniform distribution.

4.4 Uniform (0,1) distribution of F(X) where X is continuous type random variable with distribution function F(.). Application to model sampling.

4.5 Examples and problems.

Unit 5. Normal distribution. [10L,10M]

5.1 Probability density function

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) \qquad -\infty < x, \mu < \infty$$

$$\sigma > 0$$

$$Notation\ N(\mu,\sigma^2)$$

5.2 Identification of parameters μ and σ^2, nature of probability curve, symmetry of the distribution, point of inflection.

5.3 Mean, median, mode and variance.

5.4 Moment generating function, cumulant generating function, moments, recurrence relation for central moments, $\beta_1, \beta_2, \gamma_1, \gamma_2$ and standard normal distribution.

5.5 Additive property, computation of probabilities using normal probability tables.

5.6 Normal approximation to binomial and Poisson distribution, distribution of a square of a N(0,1) variate.

5.7 Examples and problems.

Unit 6. Exponential distribution. [4L,4M]

6.1 Probability density function

$$f(x)=\frac{1}{\theta}\exp(-x/\theta) \qquad x \geq 0, \theta > 0.$$

6.2 Identification of the parameter, nature of the probability curve.

6.3 Moments, moment generating function, cumulant generating function, distribution function, median and quartiles.

6.4 Lack of memory property.

6.5 Examples and problems.

Unit 7. Gamma distribution. [4L,4M]

7.1 Probability density function

$$f(x)=\frac{\alpha^{\lambda}}{\Gamma(\lambda)}e^{-\alpha x}x^{\lambda-1} \qquad x \geq 0;\ \alpha, \lambda > 0$$

$$Notation\ \ X \sim G(\alpha,\lambda)$$

7.2 Nature of the probability curve, special cases (a) $\alpha=1$, (b) $\lambda=1$.

7.3 Moment generating function, cumulant generating function, moments, cumulants, mode, γ_1, γ_2.

7.4 Distribution of the sum of n identically independently distributed exponential variates.

7.5 Additive property of gamma random variables.

7.6 Examples and problems.

Unit 8. Beta distribution of first and second kind. [6L,6M]

8.1 Probability density function of beta distribution of first kind.

$$f(x) = \frac{x^{m-1}(1-x)^{n-1}}{B(m,n)} \qquad 0 < x < 1;\ m,n > 0.$$

$$Notation\ X \sim B_1(m,n).$$

8.2 Probability density function of beta distribution of second kind.

$$f(x) = \frac{1}{B(m,n)} \frac{x^{m-1}}{(1+x)^{m+n}} \qquad 0 < x < \infty;\ m,n > 0$$

$$Notation\ X \sim B_2(m,n).$$

8.3 Relation between two kinds of variates.

8.4 Mean, variance, mode, r^{th} raw moment.

8.5 Examples and problems.

Unit 9. Chi-square distribution. [6L,6M]

9.1 Definition of χ^2 variate as sum of squares of n identically independently distributed standard normal variates.

9.2 Derivation of Probability density function of χ^2 with n degrees of freedom (using moment generating function).

Notation: $X \sim \chi^2(n)$

9.3 Nature of probability curve.

9.4 Use of χ^2 tables for calculation of the probabilities.

9.5 Mean, variance, mode, moments, γ_1, γ_2, mgf,cgf.

9.6 Normal approximations of χ^2 distribution

9.7 Fishers normal approximation of χ^2 distribution (statement only)

9.8 Examples and problems.

Unit 10. t-distribution. [6L,6M]

10.1 Definition of t statistic with n degrees of freedom in the form

$X = \frac{U\sqrt{n}}{\sqrt{V}}$ where U is N(0,1) and V is χ^2 variate with n degrees of freedom,

U and V are independent.

10.2 Derivation of Probability density function of t variate.

Notation : $X \sim t(n)$

10.3 Nature of probability curve.

10.4 Mean, variance and moments.

10.5 Statement of normal approximation to t distribution.

10.6 Use of t probability tables for calculations of probabilities.

10.7 Examples and problems.

Unit 11. F-distribution. [6L,6M]

11.1 Definition of statistic F with n_1 and n_2 degrees of freedom as

$$X = \frac{U/n_1}{V/n_2}$$ where U and V are independent χ^2 variates with n1 and n_2 degrees of freedom respectively.

11.2 Derivation of probability density function of F variate.

Notation : $X \sim F(n_1,n_2)$

11.3 Nature of probability curve.

11.4 Mean, variance, moments and mode.

11.5 Interrelations among normal, χ^2, t and F.

11.6 Use of F tables for calculations of probabilities.

11.7 Examples and problems.

Unit 12. Sampling distributions. [10L,10M]

12.1 Random sample from a continuous distribution as i.i.d.r.vs. $X_1,X_2,\ldots X_n$.

12.2 Concept of a statistic, sampling distribution of a statistic.

12.3 Sampling distribution of sample mean when a sample is selected from normal, exponential and gamma distribution.

12.4 Statement of a central limit theorem for real numbers with applications related to only tests.

12.5 Joint distribution of $\bar{X}$ *and* $S^2 = \frac{1}{n}\sum(X_i - \bar{X})^2$ for a sample from normal population using orthogonal transformation.

12.6 Examples and problems.

Unit 13. Testing of Hypothesis. [24L,20M]

13.1 Introduction, notion of hypothesis, statistical hypothesis, null hypothesis, two types of errors, critical region, critical value, level of significance, test of hypothesis, one tailed and two tailed tests, general procedure of testing of hypothesis, discussion on examples based on determination of types of errors, critical value for given test.

13.2 Large sample tests (Tests based on normal distribution).

1) $H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$, $\mu < \mu_0$, $\mu > \mu_0$.

2) $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$, $\mu_1 < \mu_2$, $\mu_1 > \mu_2$.

3) $H_0 : \Pi = \Pi_0$, $H_1 : \Pi \neq \Pi_0$, $\Pi < \Pi_0$, $\Pi > \Pi_0$.

Paper-II: Applied Statistics .

Prerequisites : The knowledge of the topics taught at F.Y.B.Sc. in Statistics.

Objectives : This paper is introduced to provide elementary introduction of the quantitative techniques involving mathematical and statistical tools. After the completion of this course, students are expected to fulfill the following objectives:

- Determination and use of multiple regression, multiple correlation and partial correlation for real life situations.
- Study of sampling methods to estimate population characteristics under various situations.
- Application of probability and probability distributions in decision making problems and quality control problems.
- Study of analysis of variance to test homogeneity of several means.
- Study of optimization techniques in distribution and assignment problems.

Unit 1. Multiple Regression (trivariate case only). [5L,5M]

1.1 Concept of multiple regression, Yules notation and fitting of regression planes by the method of least squares.

1.2 Partial regression coefficient: Definition, properties and interpretation.

1.3 Residuals: Definition, order, properties, derivation of variances and covariances.

1.4 Necessary and sufficient conditions for the coincidence of the three regression planes.

1.5 Examples and problems.

Unit 2. Multiple and Partial Correlation (trivaraite case only). [8L,8M]

2.1 Multiple correlation coefficient: Definition, derivation, interpretation and properties (i) $0 \le R_{ij.k} \le 1$, (ii) $R_{ij.k} \ge r_{ij}$

2.2 Partial correlation coefficient: Definition, derivation, interpretation and properties (i) $-1 \le r_{ij.k} \le 1$, (ii) $b_{ji.k} . b_{ij.k} = b^2_{ij.k}$

2.3 Examples and problems.

Unit 3. Sampling Methods [18L,18M]

3.1 Simple random sampling with and without replacement: Definition, inclusion probabilities. Definition of population mean, population total and population variance.

3.2 Proof of the following results:

1) Sample mean is an unbiased estimator of the population mean.

2) $N\overline{X}$ is an unbiased estimator of population total.

3) Sample mean square is an unbiased estimator of population mean Square.

4) $\text{Var}(\overline{X}) = \frac{(N-n)}{Nn} \times s^2$ and $\text{SE}(\overline{X})$ in case of SRSWOR.

5) $\text{Var}(\overline{X}) = \frac{(N-1)}{Nn} \times s^2$ and $\text{SE}(\overline{X})$ in case of SRSWR.

3.3 Simple random sampling without replacement for proportions.

Notation used: P Sample proportion, Π Population proportion.

3.4 Proof of the following results:

1) Sample proportion is an unbiased estimator of population proportion.

2) N P is an unbiased estimator of N Π.

3) $\text{Var}(P) = \frac{(N-n)}{N-1} \times \frac{\Pi(1-\Pi)}{n}$ and SE(P).

3.5 Stratified sampling: (a) stratified sample as a sample drawn from individual strata using SRSWOR in each stratum, (b) $\overline{X}_{st}$ as an estimator of population mean and $N\overline{X}_{st}$ an estimator of population total and standard error of these estimators.

3.6 Problem of allocation, proportional allocation, optimum allocation, derivation of the expression for the standard errors of the usual estimators when these allocations are used.

3.7 Gain in precision due to stratification comparison amongst SRSWOR, stratification with proportional allocation and stratification with optimum allocation.

3.8 Cost and variance analysis in stratified random sampling, minimization of variance for fixed cost. Minimization of cost for fixed variance.

3.9 Optimum allocation as a particular case of optimization in cost and variance analysis.

3.10 Examples and problems.

Unit 4. Transportation and Assignment Problem. [20L,20M]

4.1 Definition of (i) a feasible solution, (ii) a basic feasible solution and (iii) optimal solution.

4.2 Statement of transportation problem, balanced and unbalanced transportation problem.

4.3 Methods of obtaining initial basic feasible solution:

1) North west corner method.

2) Method of matrix minima (least cost method).

3) Vogel's Approximation Method (VAM).

4.4 Optimal solution of transportation problem using uv (MODI) method, uniqueness and non uniqueness of optimal solution. Degeneracy and method of resolving degenercy.

4.5 Variants in transportation problem: No allocation in a particular cell, maximization problem.

4.6 Assignment problem: Statement of assignment problem, relation to transportation problem and solution of assignment problem using Hungarian method.

4.7 Variations of the assignment problem: Unbalanced assignment problem, maximization problem, restrictions on assignments and alternate optimal solution.

4.8 Examples and problems.

Unit 5. Statistical Process Control. [24L,20M]

5.1 Introduction: Meaning and purpose of Statistical Process Control, quality of a product, need of quality control, chance and assignable causes of variation.

5.2 Control chart: Statistical basis of control chart, 3σ limits, justification of 3σ limits and criteria for detecting lack of control.

5.3 Control charts for variables: Preliminary decisions, construction of control chart for the mean and range when (i) standards are given and (ii) no standard given. Revised control limits, interpretation from the charts and determination of process mean and standard deviation from the charts. Applications of charts.

5.4 Control charts for attributes:

1) Control chart for fraction defective: Preliminary decisions, construction of control chart when (i) standards are given and (ii) no standard given. Revised control limits, interpretation from the chart and standardized control chart for variable sample size.

2) Control chart for number of defects per unit: Preliminary decisions, construction of control chart when (i) standard given and (ii) no standard given. Revised control limits, interpretation from the chart and applications of chart.

5.5 Examples and problems.

Unit 6. Simulation. [6L,6M]

6.1 Meaning of simulation.

6.2 Advantages and disadvantages of simulation.

6.3 Monte Carlo technique of simulation.

6.4 Examples and problems.

Unit 7. Statistical Decision Theory. [8L,6M]

7.1 Meaning of statistical decision theory, acts, states of nature, outcomes, pay-off and opportunity loss (or regret).

7.2 Decision making environment: Decision making under certainty, decision making under uncertainty and decision making under risk.

7.3 Decision under uncertainty: Decision rules (i) Maximax (optimistic) criterion, (ii) Maximin (pessimistic) criterion, (iii) Minimax regret criterion and Hurwicz criterion (criterion of realism).

7.4 Decision under risk: Decision rules (i) Expected value criterion and (ii) Expected opportunity loss criterion.

7.5 Construction of Decision tree.

7.6 Examples and problems.

Unit 8. Analysis of Variance. [15L,15M]

8.1 Concept of analysis of variance, chance and assignable causes of variation.

8.2 Model for one way classified data. Assumptions and interpretation.

8.3 Model for two way classified data. Assumptions and interpretation.

8.4 Concept of resolution of total sum of squares into components for one way and two way models.

8.5 Analysis of one way model: Estimation of parameters, expected values of mean sum of squares, hypothesis and its interpretation. Preparation of Analysis of Variance table.

8.6 Analysis of two way model: Estimation of parameters, expected values of mean sum of squares, hypothesis and its interpretation. Preparation of Analysis of Variance table.

8.7 Examples and problems.

Reference books

1) Fundamentals of Mathematical Statistics by S.C. Gupta and V.K. Kapoor

2) Fundamentals of Applied Statistics by S.C. Gupta and V.K. Kapoor.

3) Introduction to Statistical Quality Control by Douglas Montgomery.

4) Operations Research by Kanti Swarup, P.K. Gupta and Man Mohan.

5) Statistical Methods by G.W. Snedecor and W.G. Cochran.

Paper-III: Practicals in Statistics

Prerequisite:- The knowledge of topics taught in the theory papers

Objectives:- The main objective of this course is to prepare the students to apply the Statistical techniques learned by them to real life situations.

Note:-

(1) The total duration of practical examination shall be 5 (five) hours.

(2) The practical examination shall consist of total 100 marks. Of the 100 marks, 10 marks shall be reserved for viva-voce and 10 marks for practical journal. Thus the practical paper shall actually carry 80 marks.

(3) The examination shall be executed in two parts. Duration of each part shall be TWO & HALF hours. Each part will carry maximum 40 marks. A student will have to solve two questions out of four given questions. Each question will carry 20 marks.

(4) Each student must complete all the practicals to the satisfaction of concerned teacher.

(5) Each student must produce at the time of the practical examination, the laboratory journal of practicals completed along with the completion certificate signed by the concerned teacher and the Head of the Department.

PRACTICALS FOR PART- I

1. Fitting of Normal distribution (At least 10 classes are expected).
2. Model sampling from Normal and Exponential distribution (Parameters of the distribution should be explicitly specified).
3. Applications of Normal and Exponential distribution (Problems describing real life situations are expected).
4. Large sample tests (All the test specified in paper I are expected).
5. Test based on t distribution (All the tests specified in paper I are expected).
6. Test based on χ^2 distribution (All the tests specified in paper I are expected).
7. Test based on F distribution (All the tests specified in paper I are expected).
8. Multiple regression (Problems based on only raw data are expected).
9. Multiple and Partial correlation (Problems based on only raw data are expected).

PRACTICALS FOR PART-II

10. Analysis of variance (one way)
11. Analysis of variance (two way)
12. Simple random sampling (Estimation of mean and SE)
13. Stratified random sampling (Estimation of mean, SE, Determination of sample size using Proportional allocation, Neyman allocation, Optimum allocation).
14. Stratified random sampling (Cost and variance analysis)
15. X bar and R chart (Problems based on only raw data are expected).
16. P and C chart (Problems based on only raw data are expected).