UNIVERSITY OF PUNE
Revised Syllabus of S.Y.B.A.

(36) STATISTICS
(General and Special)

Note : (1) A student of the Three-Year B.A. Degree Course offering 'Statistics' at the special level must offer 'Mathematical Statistics' as a General level subject in all the three years of the course.

Further students of the three-year B.A. Degree Course are advised not to offer 'Statistics' as the General level unless they have offered 'Mathematical Statistics' as a General level subject in all the three years of the course.

(2) A student of three-year B.A. Degree Course offering 'Statistics' will not be allowed to offer 'Applied Statistics' in any of the three years of the course.

(3) A student offering 'Statistics' at the Special level must complete all practicals in Practical Paper to the satisfaction of the teacher concerned. He/She

(4) must produce at the time of Practical Examination, the laboratory journal along with the completion certificate signed by the Head of the Department.

(4) **Structure of evaluation of practical paper at S.Y.B.A.**

(A) Continuous internal evaluation	**Marks**
(i) Journal	10
(ii) Viva-voce	10
Total	20

(B) Annual practical examination

Section	**Nature**	**Marks**	**Time**
I	On line examination : **Note :** Question is compulsory Q1 : MSEXCEL : Execute the commands and write the same in answer book along with answers	10	Maximum 15 minutes
II	Using Calculator **Note :** Attempt any two of the following **four** questions : Q2 : Q3 : Q4 : Q5 :	60 (30 marks for each questions)	2 hours 45 minutes
III	Viva-voce	10	10 minutes
	Total B	80	3 Hours 10 minutes
	Total of A and B	100	–

(5) Duration of the practical examination be extended by 10 minutes to compensate for the loss of time for viva-voce of the candidates.

Subject : (36) (a) Statistics (General)
Title : Research Methodology, Sampling Techniques and Statistical Quality Control

First Term

1. Research Methodology: (6)

Meaning, Objectives of Research, Classification of Research.
Formulation of Research Problem and Hypotheses.
Research Design, Features of a good Research Design.

2. Measurement Scales: (6)

2.1 Identifying and deciding on the variables to be measured
2.2 Meaning of Measurement scales. Nominal scale, Ordinal scale, Interval scale, Ratio scale.
2.3 Types of Attitude Scales:
Single Item Scales :(i) Itemized Category Scales (ii) Rank Order Scales
(iii) Comparative Scales (iv) Constant Sum Scale (v) Q Sort Scales
(vi) Paired Comaprison Scales (vii) Pictorial Scales.(viii)Continuous Scales
Multi Item Scales : (i) Likert Scales (ii) Semantic Differential Scales
(iii) Thurstone Scales (iv) Associative Scales (v) Stapel Scales

3. Sample Surveys: (8)

Concepts of distinguishable elementary units, sampling frame.
Objectives of a sample survey.
Designing questionnaire, Characteristics of a good questionnaire.
Planning, execution and analysis of a sample survey. Practical problems in planning, execution and analysis of a sample survey.
Sampling and non-sampling errors with illustrations.
Study of some survey illustrating the above ideas using suitable sampling technique.

4. Sampling Techniques: (6)

Sampling from Finite Population of size N with replacement and without replacement. Population total and Mean as parameters, definitions, inclusion probabilities.

(a) Sample mean $\bar{x}$ as an estimator of population mean, derivation of its expectation and standard error.

(b) $N\bar{x}$ as an estimator of population total, derivation of its expectation and standard error.

(c) $s^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (xi - \bar{x})^2$ as an estimator of $S^2 = \frac{1}{(N-1)} \sum_{i=1}^{N} (Xi - \overline{X})^2$ an expectation of s^2.

Sampling for proportion as an application of simple random sampling with X taking value zero or one.

5. Determination of the sample size: **(4)**

Determination of the sample size for the given

(i) margin of error and confidence coefficient

(ii) coefficient of variation and confidence coefficient.

.

6. Stratified Sampling: **(10)**

(a) Stratified sampling as a sample drawn from individual strata by SRSWOR in each stratum.

(b) $\bar{x}_{st} = \frac{\sum Ni\overline{xi}}{N}$ as an estimator of population mean $\overline{X}$ and $N\bar{x}_{st}$ as an estimator of population total and standard error of these estimators.

6.2 Problem of allocation, proportional allocation, optimum allocation, derivation of the expressions for the standard errors when these allocations are used.

7. Ratio and regression methods of estimation: **(4)**

7.1 Reasoning behind using auxiliary variable in estimation.

7.2 Situations where ratio method is appropriate.

7.3 Ratio and regression estimators of the population mean and population total.

7.4 Comments regarding bias.

.

8. Systematic Sampling (Population size divisible by sample size): **(4)**

8.1 Real life situation where systematic sampling is appropriate, Techniques of drawing sample using systematic sampling.

8.2 Estimation of the population mean and population total. Standard errors of these estimators.

.

Second Term

9. Statistical Quality Control: (6)

9.1 Introduction: Meaning and purpose of SQC, Quality of a product, need of quality control, statistical quality control, process control, lot control.

9.2 Control Charts:

9.3. Chance causes and assignable causes of variation.

9.4 Statistical basis of control charts (Connection with tests of hypotheses is NOT expected).

9.5 Probability limits. 3σ limits, justification for the use of limits based on Chebychev's inequality and large sample theory.

9.6 Criteria for detecting lack of control

(i) a point outside the control limits.

(ii) Non-random variation within the control limits of the following type:

(a) A run of seven or more points above or below the control lines.

(b) Presence of trend and cycles.

(Mathematical justification is NOT expected for (ii) only).

Use of control charts for : (i) Specification (ii) Production

10. Control charts for continuous variables: (10)

10.1 Decision preparatory to control charts:

(i) choice of the variable
(ii) basis of subgroups
(iii) size of the subgroups
(iv) frequency of the subgroup.

10.2 R chart and $\overline{X}$ chart:

Purpose of R chart and $\overline{X}$ chart, construction of R chart when the process standard deviation (σ) is not given: control limits, drawing of control chart, plotting sample range values, Drawing conclusions: Determination of state of process, necessity of revision of control limits, estimate of σ ($\hat{\sigma}$). Construction of $\overline{X}$ chart when process average is not given: control limits based on $\hat{\sigma}$ drawing of control chart. Plotting sample means.

Drawing conclusion, determination of state of process, necessity of revision of control limits. Revision of control limits on $\overline{X}$ chart and R chart. Construction of R chart when the process standard deviation is specified: control limits, drawing of control chart, plotting sample range. Drawing of conclusion, decision if the process is out of control.

10.3 Construction of $\overline{X}$ chart when the process average is specified: control limits, drawing of control chart, plotting of sample means. Drawing conclusion: Determination of state of process, decision if the process is out of control.

10.4 Process Capability study: Specification limits (both or one) ,natural tolerance limits, their comparisons, decision based on these comparisons, estimate of percent defective. Shift in the process average only when process standard deviations is fixed. Evaluation of probability of catching the shift on the first sample or on the subsequent sample after the shift.

10.5 Capability indices :C_p , C_{pk} .

11. Control chart for attributes: (10)

11.1 p-chart when subgroup sizes are same and value of the process fraction defective p is specified: control limits, drawing of control chart. Plotting sample fraction defectives, Drawing conclusions: determination of state of control, interpretation of ‘high’ and ‘low’ spots, revision of control limits.

11.2 p-chart when subgroup sizes are same and value of the process fraction defective p is not specified: control limits, drawing of control chart. Plotting sample fraction defectives, Drawing conclusions: determination of state of control, interpretation of ‘high’ and ‘low’ spots, revision of control limits.

11.3 Process capability study: Shift in the process fraction defective. Evaluation of probability (using normal approximation only) of catching the shift on the first sample or on the subsequent sample after the shift.

11.4 p-chart when subgroup sizes are different and value of the process fraction defective P is not specified:

Different types of control limits:

(i) Separate control limits.
(ii) Control limits based on average sample size.
(iii) Stabilized control limits.
(iv) Control limits based on maximum and minimum sample size.

11.5 Drawing of control chart. Plotting sample fraction defective. Drawing conclusions: Determination of state of control, interpretation of 'high' and 'low' spots, revision of control limits, simple numerical problems. Identification of real life situations.

12. C-chart: (4)

12.1 Construction of c-chart when 'standard' is not given: control limits, explanation for the use of 3σ limits, drawing of control chart.

Plotting number of defects per unit. Drawing conclusions: determination of state of control, interpretation of 'high' and 'low' spot, revision of control limits, estimate of process parameter.

12.2 Construction of c-chart when standard is given control limits, justification of 3σ limits, drawing of control chart.

Plotting number of defects per unit. Drawing conclusions- determination of state of control, interpretation of 'high' and 'low' spots, revision of control limits.

13. Acceptance sampling of attributes: (18)

13.1 Concept, comparison between 100 percent inspection and sampling inspection. Procedure of acceptance sampling with rectification – single sampling plan, double sampling plan, Explanation of the terms – producer's risk, consumer's risk, AQL, LTPD, AOQ, AOQL, ASN, ATI, OC and AOQ curves.

N.B.: Distinction between type A OC curve and type B OC curve is NOT expected.

13.2 Single sampling plan.

Evaluation of Probability of acceptance using:

(i) hypergeometric (ii) binomial (iii) Poisson and (iv) normal distributions.

Statement of AOQ and ATI, Graphical determination of AOQL, Determination of a single sampling plan by lot quality and average quality approaches (numerical problems are NOT expected.

Description of Dodge and Roming tables (numerical problems are NOT expected).

13.3Double sampling plan:

Evaluation of probability of acceptance using Poisson approximation. Statement of ASN and ATI (with complete inspection of second sample). Statement of the approximate formula of AOQ. Description of Dodge and Roming tables.

13.4 Comparison of single sampling plan and double sampling plan.

******

Books Recommended

1. Cochran W.C.: Sampling Techniques,Wiley Eastern Limited.
2. Kapoor V.K. and Gupta S.C.: Fundamental of Applied Statistics, Sultan Chand and sons.
3. Daroga Singh,F.S.Chaudhary:Theory and Analysis of Sample Survey Designs,Wiley Eastern Ltd.
4.Grant E.L.:Statistical Quality Control,McGraw Hill Book Company.
5. Douglas C. Montgomery: Introduction to Statistical Quality Control, John Wiley and Sons.
6.Duncan,A.J.:Quality Control and Industrial Statistics, D.B.Taraporvala Sons and Company Pvt Ltd.
7. Ronald,E Walpole:-Introduction to Statistics, Collier Macmillan
8. Kothari C.R.: Research Methodology Methods & Techniques, New Age International Publishers
9.Panneerselvan R. :Research Methodology, Prentice-Hall of India Pvt. Ltd.
10.Santosh Gupta,:Research Methodology and Statistical Techniques, Deep and Deep Publications Pvt. Ltd.

Webpages:

1.http://www.statpages.org/
-The webpages that perform statistical calculations.
2.http://www.statsci.org/datasets.html
-Links to many datasets for Teaching & Research in Statistics.
3.http://www.stasoft.com/
-Comprehensive suit of several statistical tools for data mining.
4.http://www.sta.ucla.edu/cases
-Links to case studies in Statistics.
5.http://www.amstat.org/publications/stats
-Links to Statistical Magazines.

Subject : (36) (b) Statistics (Special)

Paper I: CONTINUOUS PROBABILITY DISTRIBUTIONS AND DEMOGRAPHY

First Term

Pre-requisite : Calculus of several variables, Maxima, Minima, Multiple integration.

1. Continuous Univariate Distributions : (12 L)

1.1 Continuous sample space : Definition, illustrations. Continuous random variable : Definition, probability density function (p.d.f.), cumulative distribution function (c.d.f.) properties of c.d.f. (without proof), probabilities of events related to r.v.

1.2 Expectation of continuous r.v., expectation of function of r.v. E[g(X)], mean, variance, geometric mean, harmonic mean, raw and central moments, skewness, kurtosis.

1.3 Moment generating function (M.G.F.) : Definition and properties, cumulant generating function, definition properties.

1.4 Mode, median, quartiles.

1.5 Probability distribution of transformation of r.v. : $Y = g(X)$. Using

(i) Jacobian of transformation for $g(\cdot)$ monotonic function and one-to-one, on to functions.

(ii) Distribution function for $Y = X^2$, $Y = |X|$ etc.

(iii) M.G.F. of g(X).

2. Continuous Bivariate Distributions (12 L)

2.1 Continuous bivariate random vector or variable (X, Y) : Joint p.d.f., joint c.d.f, properties (without proof), probabilities of events related to r.v. (events in terms of regions bounded by regular curves, circles, straight lines). Marginal and conditional distributions.

2.2 Expectation of r.v., expectation of function of r.v. E[g(X, Y)], joint moments, Cov (X, Y), Corr (X, Y), Conditional mean, conditional variance, $EE(X|Y = y) = E(X)$, regression as a conditional expectation if it is linear function of conditioning variable.

2.3 Independence of r.v. (X, Y) and its extension to k dimensional r.v. Theorems on expectation

(i) $E(X + Y) = E(X) + E(Y)$, (ii) $E(XY) = E(X) \cdot E(Y)$ if X and Y are independent, generalization to k variables. $E(aX + bY + c)$, $Var(aX + bY + c)$.

2.4 M.G.F. : $M_{X, Y}(t_1, t_2)$ properties, M.G.F. of marginal distribution of r.v.s.

properties (i) $M_{X, Y}(t_1, t_2) = M_X(t_1, 0) \cdot M_Y(0, t_2)$ if X and Y are independent r.v.s.

(ii) $M_{X+Y}(t) = M_{X, Y}(t, t)$.

(iii) $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$ if X and Y are independent r.v.s.

2.5 Probability distribution of transformation of bivariate r.v.

$U = \phi_1(X, Y)$, $V = \phi_2(X, Y)$.

3. Standard Univariate Continuous Distributions (24 L)

3.1 **Uniform or Rectangular Distribution : Notation :** $X \rightarrow U(a, b)$

p.d.f. $f(x)=\frac{1}{b-a}$, if $a \le x \le b$

$=0$, otherwise

c.d.f., sketch of p.d.f.and c.d.f., mean, variance, symmetry. Distribution of (i) $\frac{X-a}{b-a}$, (ii) $\frac{b-X}{b-a}$, (iii) $Y = F(X)$, where $F(X)$ is the c.d.f. of continuous r.v.X. Application of the result to model sampling. (Distributions of $X + Y$, $X - Y$, XY, $\frac{X}{Y}$ are not expected.)

3.2 Normal Distribution : Notation : $X \sim N(\mu, \sigma^2)$,

p.d.f. $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$, if $-\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0$

$= 0$, otherwise

p.d.f. curve identification of scale and location parameters, nature of probability curve, mean, variance, M.G.F., CG.F., central moments, cumulants, β_1, β_2, γ_1, γ_2, median, mode, quartiles, mean deviation, additive property, computations of normal probabilities using, normal probability integral tables, probability distribution of (i) $\frac{X-\mu}{\sigma}$ standard normal variable (S.N.V.), (ii) $aX + b$, (iii) $aX + bY + c$, (iv) X^2 where X and Y are independent normal variates. Probability distribution of $\bar{X}$ the mean of n i.i.d. $N(\mu, \sigma^2)$ r.v.s. Normal probability plot, q-q plot to test normality. Simulation using Box-Muller transformation. Normal approximation to (i) binomial distribution, (ii) Poisson distribution applications of normal distribution. (iii) Statement of central limit theorem for i.i.d. r.v.s. with finite positive variance.

3.3 **Exponential Distribution : Notation :** $X \sim Exp(\alpha)$,

p.d.f. $f(x) = \alpha e^{-\alpha x}$, $x \ge 0$, $\alpha > 0$

$= 0$, otherwise

nature of p.d.f. curve, interpretation of α as rate and $1/\alpha$ as mean, mean variance, M.G.F., C.G.F., c.d.f., graph of c.d.f., lack of memory property, median, quartiles. Distribution of min (X, Y) with X, Y i.i.d. exponential r.v.s.

3.4 **Gamma Distribution : Notation :** $X \sim G(\alpha, \lambda)$.

p.d.f. $f(x) = \frac{\alpha^{\lambda}}{\lceil \lambda} e^{-\alpha x} x^{\lambda-1}$, $x \geq 0,\ \alpha > 0,\ \lambda > 0$

$= 0$, otherwise

Nature of probability curve, special cases (i) $\alpha = 1$, (ii) $\lambda = 1$, M.G.F., C.G.F., moment, cumulants, β_1, β_2, γ_1, γ_2, mode, additive property. Distribution of sum of n i.i.d. exponential variables. Relation between distribution function of Poisson and Gamma variates, Recurrence relation between moments.

Second term

4. Chi-square ($\chi 2$) Distribution (8 L)

4.1 Definition of $\chi 2$ r.v. as sum of squares of i.i.d. standard normal r.v., derivation of p.d.f. of $\chi 2$ with n degrees of freedom (d.f.) using M.G.F., nature of p.d.f. curve, computations of probabilities using $\chi 2$ tables, mean, variance, M.G.F., C.G.F., central moments μ_1, μ_2 μ_3, μ_4 mode, additive property.

4.2 Normal approximation : (i) $\left(\frac{\chi 2 - n}{\sqrt{2n}}\right)$ with proof, (ii) Fisher's approximation : statement only. (Without proof)

4.3 Distribution of $\frac{\chi_1^2}{\chi_1^2 + \chi_2^2}$ and χ_1^2 / χ_2^2 where χ_1^2 and χ_2^2 are two independent chi-square r.v.s.

5. Student's t-distribution (5 L)

5.1 Definition of t r.v. with n d.f. in the form

$$t = \frac{U}{\sqrt{\chi_n^2/n}}$$

where $U \sim N(0, 1)$ and χ_n^2 is a $\chi 2$ r.v. with n d.f. and U and χ_n^2 are independent r.v.s.

5.2 Derivation of p.d.f., nature of probability curve, mean, variance, moments, mode, use of t tables for calculation of probabilities, statement of normal approximation.

6. Snedecore's F-distribution (5 L)

6.1 Definition of F r.v. with n1 and n2 d.f. as $Fn1, n2 = \dfrac{\chi^2_{n1}/n1}{\chi^2_{n2}/n2}$ where χ^2_{n1} and χ^2_{n2} are independent chi-square r.v.s with n1 and n2 d.f. respectively.

6.2 Derivation of p.d.f., nature of probabilities curve, mean, variance, moments, mode.

6.3 Distribution of 1/(Fn1,n2), use of F-tables for calculation of probabilities.

6.4 Interrelations among, $\chi 2$, t and F variaets.

7. Sampling Distributions (5 L)

7.1 Random sample from a distribution as i.i.d. r.v.s X1, X2, ..., Xn.

7.2 Notion of a statistic as function of X1, X2, ..., Xn with illustrations.

7.3 Sampling distribution of a statistic. Distribution of sample mean $\bar{X}$ from normal, exponential and gamma distribution, Notion of standard error of a statistic.

7.4 Distribution of $\dfrac{nS^2}{\sigma^2} = \dfrac{1}{\sigma^2} \sum_{i=1}^{n} (Xi - \overline{X})^2$ for a sample from a normal distribution using orthogonal transformation. Independence of $\overline{X}$ and S^2.

8. Exact Tests based on $\chi 2$, t and F distributions : (16 L)

8.1 Test for independence of two attributes arranged in 2×2 contingency table. (With Yates' correction).

8.2 Test for independence of two attributes arranged in $r \times s$ contingency table.

8.3 Test for 'Goodness of Fit'. Without rounding-off the expected frequencies.

8.4 Test for H0 : $\sigma^2 = \sigma_0^2$ against one-sided and two-sided alternatives when (i) mean is known (ii) mean is unknown.

8.5 t-tests for population mean : (i) one sample and two sample tests for one-sided and two-sided alternatives. (ii) $(1 - \alpha)$ 100% confidence interval for population mean (μ) and difference of means $(\mu_1 - \mu_2)$ of two independent normal population and confidence interval of difference of means of two independent normal populations.

8.6 Paired t-test for one-sided and two-sided alternatives.

8.7 Test for correlation coefficient H0 : $\rho = 0$, H0 : ρ i.jk = 0 against one-sided and two-sided alternatives. Using both t and F test (ANOVA).

8.8 Test for regression coefficient H0 : $\beta = 0$ against one-sided and two-sided alternatives.

8.9 Test for H0 : $\sigma_1^2 = \sigma_2^2$ against one-sided and two-sided alternatives when (i) means are known (ii) means are unknown.

9. Demorgraphy (9 L)

9.1Vital events, vital statistics, methods of obtaining vital statistics, rates of vital events, sex ratios, dependency ratio.

9.2 Death/Mortality rates : Crude death rates, specific (age, sex etc.) death rate, standardized death rate (direct and indirect), infant mortality rate.

9.3 Fertility/Birth rate : Crude birth rates, general fertility rate, specific (age, sex etc.) fertility rates, total fertility rates.

9.4 Growth/Reproduction rates : Gross reproduction rate, net reproduction rate.

9.5 Interpretations of different rates, uses and applications.

9.6 Trends in vital rates due to the latest census.

*****

S.Y.B.A.
Subject : (36) (c)Statistics (Special)
Paper II : Practical

Notes : 1. Students must complete all the practicals in each paper to the satisfaction of the teacher concerned.

2. Students must produce at the time of practical examination the laboratory journal along with the completion certificate signed by the Head of the Department.

3. Use of computer software whenever possible to be encouraged.

Preparation by Internal Examiner for Section I : Online examination :

(1) Keep at least 4 computers with latest configuration ready with battery backup and necessary software at the examination laboratory.

(2) Trivariate and bivariate data set of 10 to 20 items be fed in computer MSEXCEL spreadsheet (Trivariate data set for multiple regression plane) before the commencement of examination. Appropriate data set for time series : linear, quadratic, exponential trend fitting, exponential smoothings be entered in spreadsheet.

(3) Any other type of data required for time to time also be entered in computer spreadsheet.

Instructions to Examiners :

(1) Students are not expected to fill data items at the time of examination. They are expected to use MSEXCEL commands to operate on data set which are already fed.

(2) The question on section I are compulsory and there is no internal option.

(3) The commands of the nature attached in specimen are to be asked, so that the total marks of all asked commands will be exactly 10.

Objectives :

1. To fit various discrete and continuous probability distributions and to study various real life situations.
2. To identify the appropriate probability model, that can be used.
3. To use forecasting and data analysis techniques in case of univariate and multivariate data sets.
4. To use statistical software packages.
5. To test the hypotheses particularly about mean, variance, correlation, proportions, goodness of fit.
6. To study applications of statistics in the field of economics, demography etc.

****

Paper II Practical

Sr.No. Title of the Experiment

1. Fitting of multiple Regression planes (also using EXCEL/SPREAD SHEET)
2. Computation of multiple and partial correlation coefficients. (also using MS-EXCEL/SPREAD SHEET)
3. Fitting of Poisson distribution (also using MS-EXCEL/SPREAD SHEET).
4. Model sampling from Poisson and Geometric distributions.
5. Fitting of negative binomial distribution
6. Fitting of Normal distribution (also using MS-EXCEL/SPREAD SHEET)
7. Model sampling from normal and exponential distribution
8. Index numbers I
9. Index numbers II
10. Time series I
11. Time series II
12. Stratified sampling I
13. Stratified sampling II
14. Control charts for variables ($\overline{X}$ and R chart)
15. Control charts for attributes (P)
    (i) sample size fixed (ii) sample size variable.
16. Control chart for number of defects (C-chart)
17. Large sample tests for means and Z transformation (also using EXCEL/SPREAD SHEET)
18. Large sample tests for proportions (also using EXCEL/SPREAD SHEET)
19. Tests based on Chi-square distribution I(also using EXCEL/SPREAD SHEET)
20. Tests based on Chi-square distribution II (also using EXCEL/SPREAD SHEET)
21. Tests based on t distribution I (also using EXCEL/SPREAD SHEET)
22. Tests based on t distribution II (also using EXCEL/SPREAD SHEET)
23. Tests based on F distribution (also using EXCEL/SPREAD SHEET)
24. Selection of some problem for minor research,formulation of hypotheses,sample design,preparation of questionnaire,collection of data, analysis testing of hypotheses and conclusions.Brief report writing by the students.

Note: (1).Computer print outs are to be attached to the journal for the experiment Nos. 1,3,15 to 23

(2). Knowledge of MS-EXCEL/SPREAD SHEET should be tested on computer at the time viva-voce.

(3). Laboratory Equipments should be well equipped with sufficient number of electronic calculators and at least 4 computers with latest configuration along with necessary software, printers and UPS.

*****

S.Y.B.A.

Subject : (37) MATHEMATICAL STATISTICS (GENERAL)

Note: (1) Mathematical Statistics can be offered only as General Level Subject.

(2) A student of the Three-Year B.A. Degree course offering Mathematical Statistics will not be allowed to offer Applied Statistics in any of the three years of course.

TITLE : DISCRETE PROBABILITY DISTRIBUTIONS AND STATISTICAL METHODS

FIRST TERM

Pre-requisite : Convergence of sequences and series.

1. Discrete Probability Distribution (15 L)

1.1 Countably infinite sample space : Definition, illustrations.

1.2 Random variable (r.v.) defined on infinite sample space, probability mass function (p.m.f.), cumulative distribution function (c.d.f.), properties of c.d.f. (without proof), median, mode, probabilities of events related to random variable.

1.3 Expectation of discrete r.v., expectation of function $g(\cdot)$ of r.v., application of the result to find raw moments and central moments (upto 4^{th} order), factorial moments (upto 2^{nd} order), Properties of expectation.

1.4 Bivariate discrete random variable or vector (X, Y) taking countably infinite values : Joint p.m.f., c.d.f., properties of c.d.f. without proof, marginal and conditional distributions, independence of two r.v.s. and its extension to several r.v.s., mathematical expectation.

Theorems of expectation : (i) $E(X + Y) = E(X) + E(Y)$, (ii) $E(XY) = E(X).E(Y)$ for independent r.v.s. and generalization of these theorems to k variables. Expectation of function of random vector g(X, Y). Joint moments of order (r, s). Cov (X, Y), Corr (X, Y), Var (aX + bY + c). Conditional expectations and variance, regression coefficient using conditional expectation if it is a linear function of conditioning variable. $E\,E(X|Y = y) = E(X)$.

1.5 Moment generating function (M.G.F.) : Definition, properties, theorems on MGF. (i) $M_x(0) = 1$, (ii) M.G.F. of aX + b, (iii) M.G.F. of X + Y if X and Y are independent r.v.s. Generalization to sum of k variables, (iv) Uniqueness property (without proof).

Moments from M.G.F. using (i) expansion method, (ii) differentation method. MGF of bivariate r.v. : Definition, M.G.F. of marginal distribution of r.v.s. properties (i) $M(t_1, t_2) = M(0, t_2)\, M(t_1, 0)$ if X and Y are independent r.v.s.,

(ii) $M_{X+Y}(t) = M_{X,Y}(t, t)$, (iii) $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$ if X and Y are independent.

1.6 Cumulant generating function (C.G.F.) : Definition, properties (i) effect of change of origin, (ii) additive property. Relation between cumulants and moments (upto order four).

2. Standard Discrete Distributions (21 L)

2.1 Poisson Distribution : Notation : $X \sim P(m)$.

P.m.f. : $P(X = x)=\frac{e^{-m} m^x}{x!}$,(support) $x = 0, 1, 2 \ldots$; $m > 0$.

$=0$,elsewhere

Nature of p.m.f. Moments, M.G.F., C.G.F., mean, variance, skewness, kurtosis, mode, additive property, recurrence relation between (i) raw moments (ii) central moments, conditional distribution of X given X + Y where X and Y are independent Poisson r.v.s. real life situations.

2.2Geometric Distribution : Notation : $X \sim G(p)$,

P.m.f. : $P(X = x)=p\, q^x$,$x = 0, 1, 2, \ldots$; $0 < p < 1, q = 1 – p$

$=0$,otherwise

Nature of p.m.f. Mean, variance, distribution function, lack of memory property. Distribution of X + Y when X and Y are independent, distribution of min (X, Y) · (M.G.F., C.G.F. to be studied with negative binomial distribution). Alternative form of geometric distribution on support (1, 2, ...), real life situations.

2.3**Negative Binomial Distribution : Notation :** $X \sim> NB(k, p)$.

P.m.f. : $P(X = x)=\binom{x + k – 1}{x} p^k q^x$, $x = 0, 1, 2 \ldots.$

$0 < p < 1$, $q = 1 – p$

$=0$, otherwise

Nature of p.m.f. negative binomial distribution as a waiting time distribution. M.G.F., C.G.F., mean, variance, skewness, kurtosis (recurrence relation between moments is not expected). Relation between geometric and negative binomial distribution. Poisson approximation to negative binomial distribution. real life situations.

3. Time Series : (8 L)

3.1 Meaning and utility of time series

3.2 Components of time series; trend, seasonal variations, cyclical variations, irregular (error) fluctuations or noise

3.3 Methods of trend estimation and smoothing : (i) moving average, (ii) curve fitting by least square principle, (iii) exponential smoothing.

3.4 Measurement of of seasonal variations.

(i)simple average method, (ii) ratio to moving averages method.

3.5 Fitting of autoregressive models AR (1) and AR (2), plotting of residuals.

4. Queueing Model : (4 L)

M/M/1 : FIFO as a application of exponential distribution, Poisson distribution and Geometric distribution :

Inter-arrival rate (λ) service rate (μ), traffic intensity ($\rho = \lambda/\mu < 1$), queue discipline, probability distribution of number of customers in queue, average queue length, average waiting time in (i) queue (ii) system.

Second Term

5. Multiple Linear Regression and Multiple and Partial Correlation

(for three variables : (18 L)

5.1 Notion of multiple linear regression, Yule's notation (trivariate case – sample data only).

5.2 Fitting of regression planes by the method of least squares; obtaining normal equations, solutions of normal equations.

5.3 Residuals : Definition, order, properties, derivation of variance and covariances.

5.4 Definition and interpretation of partial regression coefficient $b_{ij.k}$, units of $b_{ij.k}$ definition of multiple correlation coefficient $R_{i.jk}$.

5.5 Derivation of the formula for the multiple correlation coefficient, also in terms of cofactors of correlation matrix.

5.6 Properties of multiple correlation coefficient (i) $0 \le R_{i.jk} \le 1$,

(ii) $R_{i.jk} \ge r_{ij}$, $R_{i.jk} \ge r_{ik}$.

5.7 Interpretation of coefficient of multiple determination $R^2_{i.jk}$ as (i) proportion of variation explained by the linear regression (ii) $R_{i.jk} = 1$, (iii) $R_{i.jk} = 0$. Adjusted $R^2_{i.jk}$. Residual plots, problem of multicolinearity introduction, introduction to stepwise regression.

5.8 Definition of partial correlation coefficient $r_{ij.k}$.

5.9 Derivation of the formula for $r_{ij.k}$, also in terms of cofactors of correlation matrix.

5.10 Properties of partial correlation coefficient (i) $-1 \le r_{i.j.k} \le 1$, (ii) $b_{ij.k}\, b_{ji.k} = r^2_{ij.k}$.

Effect of partial correlation on regression estimate.

5.11 Introduction to odds ratio and logistic regression with one regressor.

6. Tests of Hypotheses (15 L)

6.1 Statistics and parameters, statistical inference : problem of estimation and testing of hypothesis. Estimator and estimate. Unbiased estimator (definition and illustrations only). Statistical hypothesis, null and alternative hypothesis, one sided and two sided alternative hypothesis, critical region, type I error, type II error, level of significance, p-value. Confidence interval.

6.2 Tests for mean of $N(\mu, \sigma^2)$, σ known, using critical region approach

(i) $H_0 : \mu = \mu_0$ $H_1 : \mu \ne \mu_0$, $H_1 : \mu > \mu_0$, $H_1 : \mu < \mu_0$

(ii) $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \ne \mu_2$, $H_1 : \mu_1 > \mu_2$, $H_1 : \mu_1 < \mu_2$

Confidence intervals for μ and $\mu_1 - \mu_2$.

6.3 **Test Based Normal Approximation (Approximate tests) :** Using central limit theorem (using critical region approach and p value approach)

Tests for proportion, P : Parameter in binomial distribution with very large n.

(i) $H_0 : P = P_0$, $H_1 : P \neq P_0$, $H_1 : P > P_0$, $H_1 : P < P_0$

(ii) $H_0 : P_1 = P_2$, $H_1 : P_1 \neq P_2$, $H_1 : P_1 > P_2$, $H_1 : P_1 < P_2$

Confidence intervals for P and $P_1 - P_2$.

6.4 Fisher's Z transformation : Necessity of transformation :

Approximate tests : Tests for correlation coefficient of bivariate normal distribution.

(i) $H_0 : \rho = \rho_0$, $H_1 : \rho \neq \rho_0$, $H_1 : \rho > \rho_0$, $H_1 : \rho < \rho_0$

(ii) $H_0 : \rho_1 = \rho_2$, $H_1 : \rho_1 \neq \rho_2$, $H_1 : \rho_1 > \rho_2$, $H_1 : \rho_1 < \rho_2$.

7. Index Numbers (8 L)

7.1 Consumers price index numbers : Considerations in its construction : (i) family budget method (ii) aggregate expenditure method.

7.2 Shifting of base, splicing, deflating, purchasing power.

7.3 Chain base index numbers and fixed base index numbers.

8. National Income (7 L)

8.1 Definition of national income by (i) Marshall, (ii) Pigou, (iii) Fisher.

8.2 Different concept of national income (a) gross national product (GNP), (b) net national product (NNP).

8.3 Personal income, disposable income, per capita income, gross domestic product (GDP), national income at market price, national income at factor cost, national income at current prices, national income at constant prices.

8.4 Methods of estimation of national income and the difficulties in methods.

(a) output method, (b) income method, (c) expenditure method.

8.5 Importance of national income.

Books recommended :

1. Hogg, R. V. and Craig, A. T. : Introduction to Mathematical Statistics (Third Edition), Macmillan Publishing Co. Inc. 866, Third Avenue, New York 10022.
2. Gupta, S. C. and Kapoor V. K. : Fundamentals of Mathematical Statistic, Sultan Chand and Sons, 23, Daryaganj, New Delhi 110002.
3. Mood, A. M., Graybill F. A. and Boes, F. A. : (Chpaters II, IV, V, VI) Introduction to Theory of Statistics (Third Edition), McGraw – Hill Series G A 276, 1974.
4. Walpole R. E. and Mayer R. H. : Probability and Statistics (Chapter 4, 5, 6, 8, 10), Macmillan Publishing Co. Inc. 866, Third Avenue, New York 10022.
5. Arora Sanjay and Bansilal : New Mathematical Statistics : First Edition, Stya Prakashan, 16/7698 New Market, New Delhi – 5 (1989).
6. Medhi, J. : Statistical Methods, Wiley Eastern Ltd., 4835/24, Ansari Road, Daryaganj, New Delhi – 110002.
7. Mayer, P. L. : Introductory Probability and Statistical Applications, Addision Weseley Pub. Comp. London.

8. Kulkarni, M. B. Ghatpande, S. B. and Gore, S. D. : Common Statistical Tests Satyajeet Prakashan, Pune 411029 (1999).

9. Gupta, S. P. : Statistical Methods, Sultan Chand and Sons, 23, Daryaganj, New Delhi 110002.

10. Mukhopadhya Parimal (1999) : Applied Statistics, New Central Book Agency, Pvt. Ltd. Calcutta.

11. Goon A. M., Gupta, M. K. and Dasgupta, B. (1986) : Fundamentals of Statistics, Vol. 2, World Press, Calcutta.

12. Gupta, S. C. and Kapoor (V. K. (1987) : Fundamentals of Applied Statistics. S. Chand and Sons, New Delhi.

13. Sheldon Ross : A first course in probability : Pearson education publishers.

14. Neil Weiss : Introductory Statistics : Pearson education publishers.

15. M. K. Jhingan : Macro Economic Theory : Vrinda Publications Pvt. Ltd. New Delhi.

16. R. D. Gupta : Keynes Post – Keynesian Economics : Kalyani Publishers, New Delhi.

17. M. L. Sheth : Macro Economics : Lakshmi-Narayan Agarwal education publishers, Agra 3.

18. H. L. Ahuja : Modern Economics : S. Chand publishers, New Delhi.

*****

S.Y.B.A. Applied Statistics (General)

Note : (1) 'Applied Statistics' can be offered only as a General level subject.

(2) A student of Three-Year B.A. Degree course offering 'Applied Statistics' will not be allowed to offer 'Mathematical Statistics'and / or 'Statistics' in any of the three years of the course.

Subject: (38) APPLIED STATISTICS (General)

Title : Applications of Statistics and Theory of Probability.

FIRST TERM

1. Multiple regression plane, multiple and partial correlation coefficient (using tri-variate data). (9)

1.1 Notion of multiple regression plane.

1.2 Given total coefficient of correlation (rij) and standard deviation (σ) fitting of regression plane by the method of least squares and finding estimated values.

1.3 Given sums, sums of squares and sum of squares of deviations from respective mean etc. fitting og regression plane, and estimated values by the method of least squares and finding the estimated values.

1.4. Notion of multiple correlation coefficient (Ri.jk) partial correlation coefficient (rij.k) and its computations.

1.5 Simple examples and problems.

2.Index Numbers: (13)

2.1 Meaning and utility of Index Numbers, considerations arising in the construction of index numbers.Weighted and unweighted index numbers.

2.2 Various types of index numbers (viz. Laspeyres, Paache, Fisher, Walsh, Marshall Edgeworth, Dorbish-Bowley, Kelly).

2.3 Shifting of base, splicing, deflating, purchasing power.

2.4 Examples and Problems.

3 Permutations and Combinations: (5)

3.1 Definitions of permutation and combination.

3.2 Relation between permutation and combination.

(i) ${}^{n}C_{r} = {}^{n}C_{n-r}$ (ii) ${}^{n}C_{r} + {}^{n}C_{n-1} = {}^{n+1}C_{r}$ Examples and Problems.

4 Probability: (15)

4.1 Concepts of a set.

4.2 Concept and definition of union, intersection of two sets, complement of a set.

4.3 Concept of random experiment, sample space, event.

4.4 Definition of event, elementary event, certain event, impossible event, problems on sample space, events for a given random experiment.

4.5 Classical definition of probability.

4.6 Examples.

4.7 Probability model.

4.8 Axioms of probability.

4.9 Theorems of Probability (Explain through illustrations)

(i) P(A) + P(A`)=1.

(ii) $0 \leq P(A) \leq 1$.
(iii) $P(\Phi)=0$.
(iv) If $A \subseteq B$ then $P(A) \leq P(B)$.
(v) $P(AUB)=P(A)+P(B)-P(A \cap B)$.
(vi) $P(AUB) \leq P(A)+P(B)$.
(vii) Statement for 3 events for (v) and (vi)

4.10 Simple numerical problems.
4.11 Definition of conditional probability.
4.12 Deriving the formula for conditional probability.
4.13 $P(A \mid B)$ when ACB or BCA or $A \cap B=\Phi$.
4.14 Theorem on $P(A \cap B)$.
4.15 Concept and definition of independence of two events.
4.16 Pairwise independence and complete independence in case of three events.
4.17 Simple problems and examples.

5 Discrete Random Variable (r.v.) (6)

5.1 Definition of a discrete r.v.
5.2 Definition of probability mass function (p.m.f.) of a discrete r.v.
5.3 Examples.
5.4 Definition of expectation of a discrete r.v. and expectation of a linear combination of discrete r.v. X.
5.5 Definition of variance of discrete r.v. X.
5.6 Examples.

SECOND TERM

6 Special Discrete Distributions (finite sample space) (15)

6.1 Discrete uniform distribution: p.m.f. mean and variance. Illustrations of real life situations where this distribution can be applied.
6.2 Binomial distribution : Notation X~B(n,p). p.m.f., mean and variance, additive property (derivations excluded).Illustrations of real life situations where the distribution can be applied. Computation of probabilities of events related to binomial r.v..

Special Discrete Distributions (Countably infinite sample space)

6.3 Poisson distribution : Notation X~P(m) p.m.f. ,mean and variance, additive property (derivations excluded), Illustrations of real life situations where the distribution can be applied .Computation of probabilities of events related to a Poisson r.v.

6.4 Simple examples and problems.

7. Bivariate Probability Distributions: (14)

7.1 Definition of two-dimentional discrete r.v. ,its p.m.f.

7.2 Computation of probabilities of events in bivariate probability distributions.

7.3 Concepts of marginal and conditional probability distributions.

7.4 Independence of two discrete r.vs.

7.5 Definition of mathematical expectation of two dimensional discrete r.v.

7.6 Definitions of conditional mean and conditional variance.

7.7 Definition of covariance, correlation coefficient (ρ).

7.8 Examples and problems.

8. Time Series: (14)

8.1 Meaning and usefulness of time series analysis.

8.2 Components of a time series : trend , seasonal, cycle and irregular.

8.3 Additive and Multiplicative Models.

8.4Methods of estimating seasonal components:

(i) Methods of averages

(ii) Ratio to trend obtained by moving averages.

(iii) Link relative methods

(iv) Ratio to trend by least square method.

9. Elements of Demography (5)

9.1 Introduction, need of vital statistics.

9.2 Mortality Rates: Crude Death Rate(CDR),Standardized Death Rate (STDR)

9.3 Fertility and Reproduction Rates: Crude Birth Rate (CBR), General Fertility Rate(GFR).,Age-specific Fertility Rate(ASFR).Total Fertility Rate(TFR).Gross Reproduction Rate(GRR), Net Reproduction Rate(NRR).

9.4 Examples and problems.

Books Recommended

1. Gupta S.C ,Kapoor,V.K. Fundamentals of Applied Statistics,Publisher :Sultan Chand and Sonsd,New Delhi.

2. Goon,Gupta,Das Gupta,Fundamental of Statistics,Vol.II Punblisher :Shripati Bhattachrjee for the World Press Pvt. Ltd,Calcutta.

3. Lipschutz : Probability and Statistics,Publisher Schaum's Outline Series,New York.

4. Walpole,Myres:Probability and Statistics ,Publisher Mcmillan Publishing Co. New York.

5. Asthana B.N. and Srivastava S.S. : Applied Statistics of India,Published by Srivastava.

*****

Subject: (40) Statistical Pre-Requisites (Special)

The Courses in "Statistical pre-requisite" may be offered only by candidates offering one of the social sciences as their special subject at the B.A. Degree examination.

The courses in "Mathematical/Statistical pre-requisite" can not be offered by those who offered any of the courses in the Mathematics/Statistics Groups for their B.A Examination.

First Term

Probability and Probability distributions:

1 Concept of probability
2 Computation of probability by direct enumeration of cases.
3 Theorems of total and compound probability.
4 Conditional probabilities and Baye's theorem
5 Use of Difference equations in solving problems of probability.
6 Games of chance.
7 Mathematical expectation.
8 Standard distribution: Binomial Poisson. Negative Binomial. Logarithmic, hypergeomatric . Normal. Their means and variance.

Second Term

Demography:

9 Measures of Mortality
10 Construction of Life Tables
11 Mortality Projections and Theories
12 Family Formation, Composition and dissolution
13 Measures of Fertility and Reproduction..

Reference Books

1) Uspensky, J.V. : Introduction to Mathematical Probability Chs. I to V, VIII and IX
2) Kendall, M.G. and Stuarr : Advanced theory of Statistics, Vol. I, Ch. V. , Allan
3) Mortimer Siegleman : Introduction to Demography ,Chs. 4,5,6.8. and 9.

*****