UNIVERSITY OF PUNE
Revised Syllabus of S.Y.B.A. STATISTICS
(General and Special)
With Effect from June 2014

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

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

A)

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

B) Annual practical examination

Section	**Nature**	**Marks**	**Time**
I	On line examination: **Note :** Question No.1 is compulsory Q. 1: Execute the commands and write the same in answer book along with answers using MSEXCEL and R-Software	**10**	Maximum 30 minutes
II	Using Calculator **Note :** Attempt any two of the following **four** questions : Q2 : Q3 : Q4 : Q5 :	**60** (30 marks for each question)	2 hours 50 minutes
III	Viva-voce	**10**	10 minutes
	Total of (B)	**80**	3 Hours 30 minutes

	Total of (A) and (B)	**100**	

Subject : Statistics (General – I)

Title : Sampling Techniques and Statistical Process Control

First Term

1. Sampling: (12)

1.1 Simple Random Sampling from Finite Population of size(N) with replacement(SRSWR) and without replacement(SRSWOR).Population mean and Population total as parameters, inclusion probabilities.

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

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

(c) Estimator of above standard errors, in case of SRSWR and SRSWOR.

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

(a) sample proportion (p)as an estimator of population proportion(P) of units possessing a certain attribute, derivation of expectation and standard error of (p).

(b) Np as an estimator of total number of units in the population possessing a certain attribute, derivation of expectation and standard error of (Np).

(c) Estimator of above standard error both in case of SRSWR and SRSWOR.

1.4 Examples and Problems.

2. Determination of the sample size: (6)

2.1Determination of the sample size for the given:

(i) Margin of error and confidence coefficient.

(ii) Coefficient of variation and confidence coefficient.

2.2 Problems.

3. Stratified Sampling: (16)

3.1 Stratification,basis of stratification, real life situation where stratification can be used.

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

3.3 (a) $\bar{y}_{st} = \frac{N_i \bar{y}_i}{N}$ as an estimator of population mean $\bar{Y}$.Derivation of expectation and standard error of $\bar{y}_{st}$.

(b) $N\bar{y}_{st}$ as an estimator of population total , derivation of expectation and standard error of these $N\bar{y}_{st}$.

(c) Estimator of standard errors.

3.4 Derivation of allocation(n_i)under proportional allocation, optimum allocation and Neyman's allocation with the expressions for the standard errors when these allocations are used.

3.5 Examples and Problems.

4.Systematic Sampling(Population size divisible by sample size) (6)

4.1Real life situations where systematic sampling is appropriate. Techniques of drawing a sample using systematic sampling.

4.2 Estimators of the population mean and population total, standard error of these estimators**(No derivation).**

4.3 Examples and Problems.

5. Statistical Process Control (On line methods): (8)

5.1 **Introduction**

Meaning and purpose of Statistical Process Control (SPC), quality of the product, need of process control, statistical process control, on line process control methods (control charts).

5.2 **Seven Process Control (PC) Tools of SPC:**

(i) Check Sheet, (ii) Cause and effect diagram (CED),
(iii) Pareto Diagram, (iv) Histogram,
(v) Control chart, (vi)Scatter Diagram,
(vii) Design of Experiments (DOE).

5.3 Chance causes and assignable causes of variation.

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

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

5.6 **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).

Second Term

6. Control charts for variables: **(14)**

6.1 Decision preparatory to control charts:

(i) choice of the variable,

(ii) basis of subgroups,

(iii) size of the subgroups,

(iv) frequency of the subgroup.

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

(I) Purpose of R and $\overline{X}$ chart, normal probability plot for checking normality assumption. Construction of R chart when the process standard deviation is specified : control limits, drawing of control chart, plotting of sample ranges, drawing conclusion,determination of state of control process, corrective action if the process is out of control.

(II) 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 control process, corrective action if the process is out of control.

(III) Construction of R chart when the process standard deviation (σ) is not given: control limits, drawing of control chart, plotting sample range values, revision of control limits if necessary, estimate of σ (σ) for future use. Construction of $\overline{X}$ chart when the process average μ is not given : control limits based on σ, drawing of control chart, plotting sample means, revision of control limits of chart $\overline{X}$, if necessary.

Note : to find revised control limits of any control chart delete the sample points above UCL and points below LCL, in case of R and $\overline{X}$ charts, first of all, revisions of control limits of R is to be completed and then by using the observations for which R chart shows the process is under control, the control limits for $\overline{X}$ chart should be determined. Revision of control limits of $\overline{X}$ chart be continued without revising the value of *R* or σ .Estimate of μ for further use. Determination of state of control of the process. Identification of real life situations where this technique can be used. Limitations of $\overline{X}$ and R charts.

6.3 Examples and Problems.

7 Control charts for Attributes : **(14)**

7.1 **p - chart:**

Decisions preparatory to control charts

(i) Size of subgroups

(ii) Frequency of subgroups (periodicity).

Construction and working of p – chart:

(a) when subgroup sizes are same and value of the process fraction defective *p* is specified : control limits, drawing of control chart, plotting of sample fraction defectives, revision of control limits if necessary, estimation of *p* for future use. Determination of state of control of the process. Interpretation of high and low spots. Probability of detecting the shift in process fraction defective(or signal) using normal approximation.

(b) when subgroups sizes are different and value of the process fraction defective *p* is not specified (different cases of control limits):
(i) Separate control limits, (ii) control limits based on average sample size,
(iii) stabilized (standardized P)control limits, drawing of control chart, plotting sample fraction defective, determination of state of control of the process. Identification of real life situations. Limitations of p- chart.

7.2 **Confirming run length Chart (CRL)** : Concept, runs formed due to defective and good items, LCL for number of runs. Interpretation. Distinction between CLR and p- chart.

7.3 **C- chart:**

(a) Construction of C-chart when standard is given:
control limits, justification of 3 sigma limits, drawing of control chart, plotting number of defects per unit.

(b) Construction of c-chart when standard is not given:
control limits, justification for the use of 3 sigma limits, drawing of control chart. Plotting number of defects per unit, revision of control limits, if necessary, estimate of process parameter for future use. Determination of state of control, interpretation of high and low spots in above cases. Identification of real life situations. Probability of detecting shift (or signal) in parameter λ. Comparison between p and C charts. Limitations of C-chart.

7.4 Examples and Problems.

8. Statistical Process Control (Of line methods): (12)

8.1 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.

8.2 **Single sampling plan:**

Expressions of Probability of acceptance using:

(i) Hypergeometric (ii) Binomial (iii) Poisson and (iv) Normal distributions. Expressions of the formula 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).**

8.3 **Double sampling plan:**

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

8.4 Comparison of single sampling plan and double sampling plan.

8.5 Simple problems to compute probability of acceptance by using single sampling plan and double sampling plan.

8.6 Examples and Problems.

9. Capability Studies: (8)

9.1 Specification limits, natural tolerance limits and their comparisons, decisions based on these comparisons, estimate of percent defective.

9.2 Shift in the process average, evaluation of probability of detecting a shift (or getting signal) on the first sample or on the subsequent samples after the shift (when process standard deviation is fixed). Average Run Length (ARL)for $\overline{X}$ chart, Average Time to Signal (ATS). Operating Characteristic (O.C.) curve for $\overline{X}$ chart, using normality assumption.

9.3 Capability ratio and capability indices *(Cp*), capability performance indices *Cpk* with respect to machine and process, interpretation, relationship between (i) *Cp* and *Cpk* (ii) defective parts per million and *Cp.*

9.4 Examples and Problems.

Books Recommended

1. A.J.Duncan: Quality Control and Industrial Statistics, Taraporewala Sons and Co. Pvt. Ltd., Mumbai.
2. D.C. Montgomery : Statistical Quality Control, John Wiley and Sons, Inc., NewYork.
3. D.H. Besterfield, C.B. Michna etc. (3rd edition 2009) :Total Quality Management Pearson Education, Delhi.
4. D. Singh and F.S.Chaudhary : Theory and Analysis of Sample Survey Designs, Wiley Eastern Ltd., New Delhi.
5. E.L.Grant and Leavenworth : Statistical Quality Control, Mc-Graw Hill Kogakusha Ltd., New Delhi.
6. Johnson and Kotz : Capability Studies, Chapman and Hall Publishers.

7. M.N. Murthy : Sampling methods, Indian Statistical Institute, Kolkata.

8. P.V.Sukhatme and B.V.Sukhatme: Sampling theory of Surveys with Applications, Indian Society of Agricultural Statistics, New Delhi.
9. P. Mukhopadhyay : Sampling theory and methods of survey sampling.
10. W.G.Cochran: Sampling Techniques, Wiley Eastern Ltd., New Delhi.

Subject: Statistics (Special – I)
Title : Continuous Probability Distributions, Sampling Distributions and Inference

First Term

1. Continuous Univariate Distributions : (12)

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 **(C.G.F.):** definition and properties.

1.4 Mode, median, quartiles.

1.5 Probability distribution of function of r.v. : Y = g(X). **u**sing: i) Jacobian of transformation for g(.) 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).

1.6 Examples and Problems.

2. Continuous Bivariate Distributions : (12)

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) 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. : $M_{X,\,Y}$ (t1, t2), properties, M.G.F. of marginal distribution of **r. v.s.,** properties
i) $M_{X,\,Y}$ (t1, t2) = M_X (t1, 0) .M_Y (0, t2), 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) .M_Y (t) if X and Y are independent r.v.s.

2.5 Probability distribution of transformation of bivariate r. v. $U = \phi_1(X, Y)$, $V = \phi_2(X, Y)$.

2.6 Examples and Problems.

3. Standard Univariate Continuous Distributions:

3.1 **Uniform or Rectangular Distribution :** Probability density function (p. d. f.) **(03)**

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & otherwise \end{cases}$$

Notation **:** X ~ **U[a, b]**
p. d. f., sketch of p. d. f. ,c. d. f. , mean, variance, symmetry. Distribution of i) $\dfrac{X-a}{b-a}$, ii) $\dfrac{b-X}{b-a}$,, 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 and X/Y are not expected.)

3.2 **Normal Distribution :** Probability density function (p. d. f.) **(12)**

$$f(x) = \begin{cases} \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, & -\infty < x < \infty; -\infty < \mu < \infty;\ \sigma > 0 \\ 0, & otherwise \end{cases}$$

Notation: X~ N(μ, σ^2).
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. vs. Normal probability plot, q-q plot to test normality. Model sampling from Normal distribution using (i) Distribution function method & (ii)Box-Muller transformation as an application to simulation.
State & prove central limit theorem for i.i.d. r.v.s with finite positive variance.(Proof should be using m.g.f.) Its illustration for Poisson and Binomial distributions.

3.3 **Exponential Distribution :** Probability density function (p. d. f.) **(04)**

$$f(x)=\begin{cases}\alpha e^{-\alpha x} & , \quad x\geq 0,\ \alpha>0\\ 0 & , \quad otherwise\end{cases}$$

Notation **:** X ~ Exp (α).
Nature of p. d. f., density curve, interpretation of α as rate and 1/α 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 :** Probability density function (p. d. f.) **(05)**

$$f(x)=\begin{cases}\frac{\alpha^{\lambda}}{\Gamma\lambda}x^{\lambda-1}e^{-\alpha x} & , \ x\geq 0;\alpha>0\\ 0 & , \ otherwise\end{cases}$$

Notation : X~ G(α, λ).

Nature of probability curve, special cases: i) α= 1, ii) λ= 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.
3.5 Examples and Problems.

Second Term

4.Chi-square (χ^2) Distribution : (10)

4.1 Definition of χ^2 r. v. as sum of squares of i. i. d. standard normal r. vs., 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 χ^2 tables,mean, variance, M. G. F., C. G. F., central moments, β_1, β_2, γ_1, γ_2, mode, additive property.

4.2 Normal approximation: $\frac{\chi_n^2 - n}{\sqrt{2n}}$ with proof .

4.3 Distribution of $\frac{X}{X+Y}$ and $\frac{X}{Y}$, where X and Y are two independent chi-square random variables.

4.4 Examples and Problems.

5. Student's t-distribution : (06)

5.1 Definition of T r. v. with n d. f. in the form $T = \frac{U}{\sqrt{\chi_n^2 / n}}$, where U ~ 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.

5.3 Examples and Problems.

6. Snedecore's F-distribution : (06)

6.1 Definition of F r. v. with n_1 and n_2 d. f. as $F_{n_1, n_2} = \frac{\chi_{n_1}^2 / n_1}{\chi_{n_2}^2 / n_2}$ where χ_{n1}^2 and $\chi_{n_2}^2$ are independent chi-square r.v.s with n_1 and n_2 d.f. respectively .

6.2 Derivation of p. d. f., nature of **probability** 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 variates.

6.5 Examples and Problems.

7. Sampling Distributions : (08)

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 $\frac{nS^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2$ for a sample from a normal distribution using orthogonal transformation.
Independence of $\bar{X}$ and S^2.

7.5 Examples and Problems.

8. Exact Tests: **(18)**

8.1 Tests based on chi-square distribution:

a) Test for independence of two attributes arranged in 2 x2 contingency table. (With Yates' correction).

b) Test for independence of two attributes arranged in r x s contingency table, McNemar's test

c) Test for 'Goodness of Fit'. **(Without rounding-off the expected frequencies).**

d) Test for $H_0 : \sigma^2 = \sigma_0^2$ against one-sided and two-sided alternatives when i) mean is known , ii) mean is unknown.

8.2 Test**s** based on t-distribution:

a) t-tests for population means : i) one sample and two sample tests for one-sided and two-sided alternatives, ii) (1 –α) 100% two sided confidence interval for population mean (μ) and difference of means (μ1 – μ2) of two independent normal population**s** and confidence interval of difference of means of two independent normal populations.

b) Paired t-test for one-sided and two-sided alternatives.

8.3 Test based on F-distribution:

Test for $H_0 : \sigma_1^2 = \sigma_2^2$ against one-sided and two-sided alternatives when i) means are known, ii) means are unknown.

8.4 Examples and Problems.

Subject: Statistics (Special –II)

Title: Practical

Notes : 1. Students must complete all the practicals 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.

Sr. No.	**Title of the experiment**	**No. of practicals**
1	Fitting negative binomial distribution, testing goodness of fit.	**1**
2	Fitting of normal distribution, testing goodness of fit (also using qq-plot).	**1**
3	Model sampling from exponential, normal distribution using (i) distribution function, (ii) Box-Muller transformation.	**1**
4	Time series : Estimation and forecasting of trend by fitting of AR (1) model, exponential smoothing, moving averages.	**1**
5	Estimation of seasonal indices by ratio to trend	**1**
6	Test for means and construction of confidence interval . Also using MSEXCEL (i) $H_0 : \mu = \mu_0$, σ known and σ unknown (ii) $H_0 : \mu_1 = \mu_2$, σ_1, σ_2 known (iii) $H_0 : \mu_1 = \mu_2$, $\sigma_1 = \sigma_2 = \sigma$ unknown (iv) $H_0 : \mu_1 = \mu_2$, paired t test	**2**
7	Tests based on χ^2 distribution (i) goodness of fit (ii) independence of attributes (2 x2, m x n contingency table), (iii)Mc Nemar's test, (iv) $H_0 : \sigma^2 = \sigma_0^2$, μ unknown, confidence interval for σ^2	**2**
8	Tests based of F distribution $H_0 : \sigma_1^2 = \sigma_2^2$ (i)mean known, (ii) mean unknown	**1**
9	Stratified sampling	**2**
10	Control charts for variables ($\overline{X}$ and R chart)	**1**
11	Control charts for attributes (p chart)	**1**

	(i) sample size fixed (ii) sample size variable.	
12	Fitting of multiple regression plane and copmputation of multiple and partial correlation coefficients using MSEXCEL.	**1**
13	Fitting of normal distribution using MSEXCEL.	**1**
14	Computations of probabilities of Normal, Exponential gamma, $\chi 2$, t, F using R.	**1**
15	Use of basic R software commands, finding summary statistics using R software	**1**
16	Tests using R software	**1**
17	**Project :** Practicals based on analysis of data collected by students in a batch of size not exceeding 6 students.	**5**

Note: (1).Computer print outs are to be attached to the journal for the experiment Nos. 6,12,13,14,15 and 16.

(2). Knowledge of MS-EXCEL/SPREAD SHEET and R 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.

Subject: Mathematical Statistics (General-II)

Title: Discrete Probability Distributions , Statistical Methods and R- Software

First Term

1. Standard Discrete Distributions:

1.1 **Negative Binomial Distribution :** Probability mass function (p. m. f.) **(06)**

$$P(X=x)=\begin{cases}\binom{x+k-1}{x}p^k\,q^x\,, & x=0,1,2,\cdots\cdots;0<p<1;q=1-p\\ 0\,, & otherwise\end{cases}$$

Notation **:** X ~ NB (k, p).

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.

1.2 **Multinomial Distribution :** Probability mass function (p. m. f.) **(12)**

$$P(X_1 = x_1, X_2 = x_2, \cdots, X_k = x_k) = \begin{cases} \dfrac{n!\, p_1^{x_1}\, p_2^{x_2} \cdots p_k^{x_k}}{x_1!\, x_2! \cdots x_k!}, & x_i = 0,1,2,\cdots n,\ i = 1,2,\cdots,k; \\ & x_1 + x_2 + \cdots + x_k = n; \\ & 0 < p_i < 1;\ i = 1,2,\cdots,k \\ & p_1 + p_2 + \cdots + p_k = 1; \\ 0, & \textit{otherwise} \end{cases}$$

Notation : $(X_1, X_2, \cdots, X_k) \sim MD(n, p_1, p_2, \cdots p_k)$, $\underline{X} \sim MD(n, \underline{p})$

where $\underline{X} = (X_1, X_2, \cdots, X_k)$, $\underline{p} = (p_1, p_2, \cdots p_k)$.

Joint MGF of (X1, X2 , …. ,Xk), use of MGF to obtain means, variances, covariances, total correlation coefficients, multiple and partial correlation coefficients for k= 3, univariate marginal distribution, distribution of Xi+Xj. Conditional distribution of Xi given Xi+Xj= r, Variance – covariance matrix, rank of variance – covariance matrix and its interpretation and Real life situations and applications.

1.3 Truncated Distributions (06)

Concept of Truncated distribution truncation to the right, left and on both sides. Binomial distribution B(n, p) left truncated at X=0, (value zero is discarded), its p.m.f. mean, variance

Poisson distribution P(m), left truncated at X=0 ,(value zero is discarded),its p.m.f. mean, variance.

Real life situations and applications.

2. Time Series : (18)

2.1 Meaning and utility of time series, Components of time series**:** trend, seasonal variations, cyclical variations, irregular (error) fluctuations or noise.

2.2 Exploratory data analysis: Time series plot to (i) check any trend , seasonality in the time series (ii) learn how to capture trend.

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

2.4 Measurement of seasonal variations **:** i) simple average method, ii) ratio to moving average method, iii)) ratio to trend where trend is calculated by method of least squares.

2.5 Choosing parameters for smoothing and forecasting.

2.6 Forecasting based on exponential smoothing.

2.7 Double exponential smoothing i.e. Hold Winter method

2.8 Fitting of autoregressive models AR (1), plotting of residuals.

2.9 Data Analysis of Real Life Time Series:

Price index series, share price series, economic time series, Sales tax series, market price of daily consumables, Weather related time series: Temperature and rainfall time series, wind speed time series, pollution levels.

Examples and Problems.

3. Fundamentals of R-Software : (6)

3.1Introduction to R, features of R, starting and ending R session, getting help in R, R commands and case sensitivity

3.2 Vectors and vector arithmetic

a)creation of vectors using functions c, **,**seq, rep

b)Arithmetic operations on vectors using operators +, -, *, /, ^.

c)Numerical functions: log10,log,sort,max,min, unique, range, length, var, prod, sum, summary, fivenum etc.

d) accessing vectors

3.3 Data frames : creation using data frame, subset and transform commands.

3.4 Resident data sets : Accession and summary

3.5 p,q,d,r, functions.

Second Term

4. Multiple Linear Regression : (12)

4.1Notion of multiple linear regression, Yule's notation (trivariate case).

4.2Fitting of regression plane of Y on X_1 and X_2, $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$ by the method of least squares; obtaining normal equations, solutions of normal equations.

4.3Residuals : Definition, order, derivation of variance, properties.

4.4Definition and interpretation of partial regression coefficient**s** $\beta_1 (b_{YX_1.X_2})$ and β_2 $(b_{YX_2.X_1})$. Derivation of the expression for the partial regression coefficient.

4.5Definition of multiple correlation coefficient $R_{Y.X_1X_2}$. Derivation of the expression for the multiple correlation coefficient. Properties of multiple correlation coefficient

i) $0 \le R_{Y.X_1X_2} \le 1$, ii) $R_{Y.X_1X_2} \ge \min\{r_{YX_1}, r_{YX_2}\}$.

4.6Definition of coefficient of multiple determination $R^2_{Y.X_1X_2}$ Interpretation as i) proportion of variation explained by the linear regression ii) $R^2_{Y.X_1X_2} = 1$, iii) $R^2_{Y.X_1X_2} = 0$.

4.7Definition of partial correlation coefficient $r_{YX_1.X_2}$ and $r_{YX_2.X_1}$. Derivation of the expression for the partial correlation coefficient

4.8Properties of partial correlation coefficient: i) $-1 \le r_{YX_1.X_2} \le 1$ and ii) $-1 \le r_{YX_2.X_1} \le 1$.

4.9 Examples and Problems.

5.Tests of Hypotheses : (18)

5.1Statistics 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, power of the test, level of significance, p-value. Two sided confidence

interval, finding probabilities of type I error and type II error when critical regions are specified .

5.2 Tests for mean of N (μ, σ^2 **),σ** known, using critical region approach

i) H0 : $\mu = \mu_0$ against H1 : $\mu \neq \mu_0$, H1 : $\mu > \mu_0$, H1 : $\mu < \mu_0$,

ii) H0 : $\mu_1 = \mu_2$ against H1 : $\mu_1 \neq \mu_2$, H1 : $\mu_1 > \mu_2$, H1 : $\mu_1 < \mu_2$.

Two sided confidence intervals for μ and $\mu_1 - \mu_2$.

5.3Tests Based on Normal Approximation **:** Using central limit theorem (using critical region approach and p value approach) .

Tests for population proportion P :

i) H0 : P = P0 against H1 : P ≠ P0, H1 : P > P0, H1 : P < P0

ii) H0 : P1 = P2 against H1 : P1 ≠ P2, H1 : P1 > P2, H1 : P1< P2 .

Two sided confidence intervals for P and P1 – P2.

5.4 Examples and Problems.

6. Tests of hypothesis using R-Software: (6)

6.1 Drawing a sample from population using SRSWR, SRSWOR.

6.2Tests: Z test, t test, F test and tests for proportions.

6.3 Examples and Problems.

7.Demography : (6)

7.1Vital events, vital statistics, methods of obtaining vital statistics, rates of vital events, sex ratios, dependency ratio.

7.2Death/Mortality rates : Crude death rates, specific (age, sex etc.) death rate, standardized death rate (direct and indirect), infant mortality rate.

7.3Fertility/Birth rate : Crude birth rates, general fertility rate, specific (age, sex etc.) fertility rates, total fertility rates.

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

7.5 Interpretations of different rates, uses and applications.

7.6 Trends in vital rates due to the latest census.

7.8 Examples and Problems.

8.Queueing Model : (6)

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. Examples and Problems.

Books recommended :

1. A.Sanjay and Bansilal(1989) : New Mathematical Statistics(First Edition), Satya Prakashan, 16/7698 New Market, New Delhi – 5.
2. A.M.Goon,M.K. Gupta and B. Dasgupta(1986) : Fundamentals of Statistics,Vol. 2, World Press, Calcutta.
3. A.M.Mood,F.A. Graybill and F.A.Boes(1974) : Introduction to Theory of Statistics (Third Edition), McGraw – Hill Series G A 276.

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

5. J.Medhi : Statistical Methods, Wiley Eastern Ltd., 4835/24, Ansari Road, Daryaganj, New Delhi – 110002.
6. M. K. Jhingan : Macro Economic Theory : Vrinda Publications Pvt. Ltd. New Delhi.

7. M.B.Kulkarni ,S. B. Ghatpande, S. B. and S.D.Gore(1999) : Common Statistical Tests Satyajeet Prakashan, Pune 411029.
8. M. L. Sheth : Macro Economics : Lakshmi-Narayan Agarwal education publishers, Agra 3.

9. N.Weiss : Introductory Statistics : Pearson education publishers.
10. P.L.Mayer: Introductory Probability and Statistical Applications, Addision Weseley Pub. Comp. London.
11. P. Mukhopadhya (1999) : Applied Statistics, New Central Book Agency, Pvt. Ltd. Calcutta.
12. R.V.Hogg and A.T. Craig: Introduction to Mathematical Statistics (ThirdEdition), Macmillan Publishing Co. Inc. 866, Third Avenue, New York 10022.
13. R.Sheldon : A first course in probability : Pearson education publishers.

14. R. D. Gupta : Keynes Post – Keynesian Economics :Kalyani Publishers,New Delhi.
15. R.E. Walpole and R.H.Mayer : Probability and Statistics , Macmillan Publishing Co. Inc. 866, Third Avenue, New York 10022.
16. S.C.Gupta and V.K. Kapoor : Fundamentals of Mathematical Statistic, Sultan Chand and Sons, 23, Daryaganj, New Delhi 110002.

17. S.P.Gupta : Statistical Methods, Sultan Chand and Sons, 23, Daryaganj, New Delhi 110002.
18. S.C. Gupta and V.K.Kapoor (1987) : Fundamentals of Applied Statistics. S.Chand and Sons, New Delhi.

S.Y.B.A.
Applied Statistics

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' , Statistical Pre-requisites and / or 'Statistics' in any of the three years of the course.

Subject: Applied Statistics
Title: Applications of Statistics and Theory of Probability

FIRST TERM

1. Multiple regression plane, multiple and partial correlation coefficient (using tri-variate data): (10)

(No proof is required for the derivation of the equation of multiple regression plane)

1.1 Notion of multiple regression.
1.2 Equation of Multiple regression model : $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$
1.3 Obtaining the equation of plane of regression given:
- the means, standard deviations and total correlation coefficients and obtaining the estimated value.
- sums, sums of squares and sum of squares of deviations from respective mean etc. and obtaining the estimated value.

1.4 Notion of multiple and partial correlation .
1.5 Definition of multiple correlation coefficient ($R_{i.jk}$) and partial correlation coefficient ($r_{ij.k}$)
1.6 Examples and problems.

2. Time Series: (16)

2.1 Meaning of Time Series
2.2 Various components of a time series (Explanation and illustrations of each component)
2.3 Additive and multiplicative models of time series.
2.4 Meaning and usefulness of time series analysis.
2.4 Methods of estimating trends:
(i) Freehand or graphical method (ii) Method of least square(line and second degree curve) (iii) Method of semi-averages (iv) Method of moving averages.
2.5Methods of estimating seasonal components:
(i) Methods of simple averages.
(ii) Ratio to trend obtained by moving averages.
2.6 Autoregression model, Fitting of AR(1)model.
2.7 Simple examples and problems.

3. Permutations and Combinations: (6)

3.1 Definitions of permutation and combination.
3.2 Relation between permutation and combination. (Explain through illustrations)
(i) $^{n}C_{r} = {}^{n}C_{n-r}$
(ii) $^{n}C_{r} + {}^{n}C_{r-1} = {}^{n+1}C_{r}$
3.3 Examples and Problems.

4. Probability : (16)

4.1 Revision of set theory.
Concept and definition of union, intersection of two sets, complement of a set. (no problem should be asked on this topic in the examination)
4.2 Concept of random experiment, sample space with its types.
4.3 Event and types of event: complementary event, elementary event, certain event, impossible event, mutually exclusive events and exhaustive events.
4.4 Classical definition of probability and its limitations.
4.5 Probability model.
4.6 Axioms of probability.
4.7 Theorems of Probability (Without proof explain only through illustrations)
(i) P(A) + P(A`)=1.

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

4.8 Definition of conditional probability.
4.9 Particular cases of $P(A \mid B)$ when $A \subset B$ or $B \subset A$ or $A \cap B=\Phi$.
4.10 Multiplication theorem on $P(A \cap B)$.
4.11 Concept and definition of independence of two events.
4.12 Pair-wise independence and complete independence in case of three events.
4.13 Examples and problems.

SECOND TERM

5. Univariate discrete probability distributions: (8)

5.1 Definition of a discrete random variable (r.v.)
5.2 Definition of probability mass function (p.m.f.) of a discrete r.v, cumulative distribution function(c.d.f.) and its properties.
5.3 Definition of expectation of a discrete r.v. and expectation of a linear function of discrete r.v.
5.4 Definition of variance of discrete r.v.
5.5 Examples and Problems.

6. Bivariate Probability Distributions: (13)

6.1 Definition of two-dimentional discrete r.v. ,its p.m.f.
6.2 Computation of probabilities of events in bivariate probability distributions.
6.3 Concepts of marginal and conditional probability distributions.
6.4 Independence of two discrete r.vs.
6.5 Definition of mathematical expectation of two dimensional discrete r.v.
6.6 Definitions of conditional mean and conditional variance.
6.7 Definition of covariance and correlation coefficient (ρ).
6.8 Examples and problems.

7. Special Discrete Probability Distributions : (15)

7.1 Discrete uniform distribution: p.m.f. mean and variance. Illustrations of real life situations where this distribution can be applied.
7.2 Binomial distribution : Notation $X \sim B(n,p)$. 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.
7.3 Poisson distribution : Notation $X \sim P(m)$ 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.
7.4 Examples and problems.

8. Elements of Demography : (12)

8.1 Introduction to demography, need of vital statistics and methods of obtaining vital statistics.
8.2 Mortality Rates: Crude Death Rate(CDR),Age-specific death rates(ASDR), Standardized Death Rate(STDR) (Direct and indirect method).

8.3 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).

8.4 Examples and problems.

Books Recommended

1. G. Gupta and D. Gupta: Fundamental of Statistics, Vol.II , Shripati Bhattachrjee for the World Press Pvt. Ltd, Calcutta.
2. J.V.Uspensky: Introduction to Mathematical Probability.
3. Lipschutz : Probability and Statistics, Schaum's Outline Series,New York.
4. M.G.Kendall and Stuarr : Advanced theory of Statistics, Vol. I, Allan.
5. M. Siegleman : Introduction to Demography .
6. S.C.Gupta and V.K.Kapoor: Fundamentals of Applied Statistics, Sultan Chand and Sons,New Delhi.
7. Walpole and Myres :Probability and Statistics , Mcmillan Publishing Co. New York.

S.Y.B.A.
Statistical Pre-requisites

Note : (1) Statistical Pre-requisites can be offered only as a General level subject.
(2) A student of Three-Year B.A. Degree course offering Statistical Pre-requisites will not be allowed to offer 'Mathematical Statistics' ,'Statistics' and / or **'**Applied Statistics' in any of the three years of the course.

Subject: Statistical Pre-requisites
Title: Applications of Statistics and Theory of Probability

FIRST TERM

1. Multiple regression plane, multiple and partial correlation coefficient (using tri-variate data): (10)

(No proof is required for the derivation of the equation of multiple regression plane)

1.1 Notion of multiple regression.

1.2 Equation of Multiple regression model : $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$

1.3 Obtaining the equation of plane of regression given:

- the means, standard deviations and total correlation coefficients and obtaining the estimated value.
- sums, sums of squares and sum of squares of deviations from respective mean etc. and obtaining the estimated value.

1.4 Notion of multiple and partial correlation .

1.5 Definition of multiple correlation coefficient ($R_{i.jk}$) and partial correlation coefficient ($r_{ij.k}$)

1.6 Examples and problems.

2. Time Series: (16)

2.1 Meaning of Time Series

2.2 Various components of a time series (Explanation and illustrations of each component)

2.3 Additive and multiplicative models of time series.

2.4 Meaning and usefulness of time series analysis.

2.4 Methods of estimating trends:

(i) Freehand or graphical method (ii) Method of least square(line and second degree curve) (iii) Method of semi-averages (iv) Method of moving averages.

2.5Methods of estimating seasonal components:

(i) Methods of simple averages.

(ii) Ratio to trend obtained by moving averages.

2.6 Autoregression model, Fitting of AR(1)model.

2.7 Simple examples and problems.

3. Permutations and Combinations: (6)

3.1 Definitions of permutation and combination.

3.2 Relation between permutation and combination. (Explain through illustrations)

(i) ${}^{n}C_{r} = {}^{n}C_{n-r}$

(ii) ${}^{n}C_{r} + {}^{n}C_{r-1} = {}^{n+1}C_{r}$

3.3 Examples and Problems.

4. Probability : (16)

4.1 Revision of set theory.

Concept and definition of union, intersection of two sets, complement of a set. (no problem should be asked on this topic in the examination)

4.2 Concept of random experiment, sample space with its types.

4.3 Event and types of event: complementary event, elementary event, certain event, impossible event, mutually exclusive events and exhaustive events.

4.4 Classical definition of probability and its limitations.

4.5 Probability model.

4.6 Axioms of probability.

4.7 Theorems of Probability (Without proof explain only through illustrations)

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

(ii) $0 \leq P(A) \leq 1$.

(iii) $P(\Phi)=0$.

(iv) If $A \subset B$ 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).

4.8 Definition of conditional probability.

4.9 Particular cases of $P(A \mid B)$ when $A \subset B$ or $B \subset A$ or $A \cap B=\Phi$.

4.10 Multiplication theorem on $P(A \cap B)$.

4.11 Concept and definition of independence of two events.

4.12 Pair-wise independence and complete independence in case of three events.

4.13 Examples and problems.

SECOND TERM

5. Univariate discrete probability distributions: (8)

5.1 Definition of a discrete random variable (r.v.)

5.2 Definition of probability mass function (p.m.f.) of a discrete r.v, cumulative distribution function(c.d.f.) and its properties.

5.3 Definition of expectation of a discrete r.v. and expectation of a linear function of discrete r.v.

5.4 Definition of variance of discrete r.v.

5.5 Examples and Problems.

6. Bivariate Probability Distributions: (13)

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

6.2 Computation of probabilities of events in bivariate probability distributions.

6.3 Concepts of marginal and conditional probability distributions.

6.4 Independence of two discrete r.vs.

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

6.6 Definitions of conditional mean and conditional variance.

6.7 Definition of covariance and correlation coefficient (ρ).

6.8 Examples and problems.

7. Special Discrete Probability Distributions : (15)

7.1 Discrete uniform distribution: p.m.f. mean and variance. Illustrations of real life situations where this distribution can be applied.

7.2 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.

7.3 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.

7.4 Examples and problems.

8. Elements of Demography : (12)

8.1 Introduction to demography, need of vital statistics and methods of obtaining vital statistics.

8.2 Mortality Rates: Crude Death Rate(CDR),Age-specific death rates(ASDR), Standardized Death Rate(STDR) (Direct and indirect method).

8.3 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).

8.4 Examples and problems.

Books Recommended

1. G. Gupta and D. Gupta: Fundamental of Statistics, Vol.II , Shripati Bhattachrjee for the World Press Pvt. Ltd, Calcutta.
2. J.V.Uspensky: Introduction to Mathematical Probability.
3. Lipschutz : Probability and Statistics, Schaum's Outline Series,New York.
4. M.G.Kendall and Stuarr : Advanced theory of Statistics, Vol. I, Allan.
5. M. Siegleman : Introduction to Demography .

6. S.C.Gupta and V.K.Kapoor: Fundamentals of Applied Statistics, Sultan Chand and Sons,New Delhi.
7. Walpole and Myres :Probability and Statistics , Mcmillan Publishing Co. New York.

Table with code and Title

Code	Group	2008 Pattern		2013 Pattern	
		Subject	**Title**	**Subject**	**Title**
2384	L	Statistics (GEN. – I)	Research Methodology, Sampling Techniques and Statistical Quality control	Statistics (GEN. – I)	Sampling Techniques and Statistical Process control
2385	L	Statistics (SPL. – I)	Continuous Probability Distributions and Demography	Statistics (SPL. – I)	Continuous Probability Distributions, Sampling Distributions and Inference
2386	--	Statistics (SPL. – II)	Practical	Statistics (SPL. – II)	Practical
2324	J	Mathematical Statistics (GEN.II)	Discrete Probability Distributions and Statistical Methods	Mathematical Statistics (GEN.II)	Discrete Probability Distributions , Statistical Methods and R-Software
2354	K	Statistical Pre-requisites	--------	Statistical Pre-requisites	Applications of Statistics and Theory of Probability
2414	L	Applied Statistics	Applications of Statistics and Theory of Probability	Applied Statistics	Applications of Statistics and Theory of Probability

Important Note: Keep **Statistics (General), Applied Statistics and Statistical Pre-requisites** in **one group(i.e L)** which is more appropriate. So that their examination will be on same day and same time.