Savitribai Phule Pune University

(Formerly University of Pune)

REVISED SYLLABUS

OF

S. Y. B. Sc. STATISTICS

(at subsidiary level)

Choice Based Credit System Syllabus (pattern 2019-20)

With Effect from June 2020

STATISTICS

Notes :

1. A student of the three year $B.Sc.$ degree course will not be allowed to offer Statistics and Statistical Techniques simultaneously in any of the three years of the course.
2. Students offering Statistics at the first year of the three year $B.Sc.$ course may be allowed to offer Statistical Techniques as one of their subjects in the second year of the three year $B.Sc.$ course in the place of Statistics.
3. Students offering Statistical Techniques at the first of the three year $B.Sc.$ course may be allowed to offer Statistics as one of their subjects in the second year of the three year $B.Sc.$ course in place of Statistical Techniques provided they satisfy other requirements regarding subject combinations, if any.
4. Students must complete all the practicals to the satisfaction of the teacher concerned.
5. At the time of practical examination, a student must produce the laboratory journal along with the completion certificate signed by the Head of the Department.
6. **Study Tour**: In order to acquaint the students with applications of Statistical methods in various fields such as industries, agricultural sectors , Government Institutes etc. study tour may be arranged and the report be attached in practical journal.
7. Structure of evaluation of practical paper at $S.Y.B.Sc.$

A) Continuous Internal Assessment (CIA)

Section	Marks
i) Journal	05
ii) Viva-voce	05
iii) Project	05
Total of A	**15**

B) End of Semester Examination (ESE)

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 (A) $Ms - EXCEL$ (For Sem-III) (B) $R - Software$ (For Sem-IV)	 07 07	Maximum 30 minutes
II	Using Calculator / Computer Note : Attempt any two of the four questions : Q2 : Q3 : Q4 : Q5 :	24 (12 marks for each question)	2 hours - for calculator user 1 hour 30 minutes- for computer user
III	Viva-voce	04	10 minutes
	Total of B	**35**	**2 Hours 40 minutes- for calculator user** **2 hour 10 minutes- for computer user**
	Total of A and B	**50**	

Preparation by Internal Examiner for Section I (Online examination) :

1. Keep at least 15 **computers** with latest configuration ready with battery backup and necessary software, printers, scientific calculators, necessary statistical tables, normal probability paper at the examination laboratory.
2. Appropriate data sets for time series: linear, quadratic, exponential trend fitting, exponential smoothing be entered in spreadsheet.
3. Any other type of data required as per slip 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 $Ms-EXCEL$ and R −commands (whichever is applicable) to operate on the data set which are already fed.
2. The questions on section I (On line examination Using $Ms-EXCEL$ / R−commands (whichever is applicable)) are compulsory and there is no internal option.
3. The slips made available for Section I shall be allotted to the candiadtes at random so that the total marks of all asked commands will be exactly 07.

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 various hypotheses of significance like means, proportions, independence of attributes, variance etc. included in theory (using calculators, software).
6. To compute probabilities of discrete and continuous probability distributions using MS-Excel and/or R software (whichever is applicable).
7. To study applications of statistics in the field of demography etc.

Instruction for Examination

1. The theory question paper for each paper shall cover all the topics in the pertaining syllabus with proportional weightage to the number of hours of instruction prescribed.
2. The practicals are to be conducted in batches as per the University norms for the faculty of science.
3. Medium of Instruction: English
4. Examination:

A) Pattern of examination: Semester wise

B) Standard of passing : As per norms of University

S.Y.B.Sc. STATISTICS Syllabus

For Choice Based Credit System (pattern 2019-20)

To be implemented from the Academic year 2020-2021

Structure of the course:

	Semester- III		Credits	Semester-IV		Credits	Lecture/ Week	Marks Internal | External	
Paper-I	ST-231	Discrete Probability Distributions And Time Series	02	ST-241	Tests of Significance And Statistical Methods	02	03 (each lecture of 50 minutes)	15	35
Paper-II	ST-232	Continuous Probability Distributions	02	ST-242	Sampling Distributions And Exact tests	02	03 (each lecture of 50 minutes)	15	35
Paper-III	ST-233	Statistics Practical	02	ST-243	Statistics Practical	02	Each practical is of duration 04 hours.& 20 minutes.	15	35

Equivalence for courses (2014-15 pattern) with new Course (2019-20 pattern):

Semester- III		Semester-IV	
Old Course (2014-15 pattern)	New Course (2019-20 pattern)	Old Course (2014-15 pattern)	New Course (2019-20 pattern)
ST-211: Discrete Probability Distributions, Time Series & R Software	ST-231: Discrete Probability Distributions And Time Series	ST-221: Statistical Methods and Use of R Software	ST-241:Tests of Significance And Statistical Methods
ST-212: Continuous Probability Distributions	ST-232: Continuous Probability Distributions	ST-222: Sampling Distributions and Inference	ST-242: Sampling Distributions And exact tests
ST-223: Statistics Practical(Annual Examination)	**No equivalence**	**----------**	**----------**

SEMESTER – III

PAPER - I

ST – 231: DISCRETE PROBABILITY DISTRIBUTIONS AND TIME SERIES

1. Negative Binomial Distribution: **(07 L)**

Probability mass function (p.m.f.)

$P(X = x) = \binom{x+k-1}{x} p^k q^x \quad ; \quad x = 0,1,2,\ldots \; ; \; 0 < p < 1 \; ; \; q = 1 - p \; ; \; k > 0$

$= \quad 0 \quad ; \quad otherwise.$

Notation: $X \sim NB(k,p)$.

Graphical nature of p.m.f., negative binomial distribution as a waiting time distribution, moment generating function(MGF), cumulant generating function(CGF), mean, variance, skewness, kurtosis(recurrence relation between moments is not expected), additive property of NB(k,p). Relation between geometric distribution and negative binomial distribution. Poisson approximation to negative binomial distribution. Real life situations.

2. Multinomial Distribution: Probability mass function (p.m.f.) **(10 L)**

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \frac{n!\, {p_1}^{x_1} {p_2}^{x_2} \dots {p_k}^{x_k}}{x_1!\, x_2! \dots x_k!} \quad ; \quad x_i = 0, 1, 2, \dots, n - \sum_{1}^{i-1} x_r,$$

$$i = 1,2, \dots, k$$

$$x_1 + x_2 + \cdots + x_k = n;$$

$$0 < p_i < 1;\ i = 1, 2, \dots, k;$$

$$p_1 + p_2 + \cdots + p_k = 1;$$

$$= 0 \quad ; otherwise.$$

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

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

Joint MGF of $(X_1, X_2, \dots, X_k)$, use of MGF to obtain means, variances, covariances, total correlation coefficients, variance – covariance matrix, rank of variance – covariance matrix and its interpretation, additive property of multinomial distribution, univariate marginal distribution, distribution of $Xi + Xj$, conditional distribution of Xi given $Xj = r$, conditional distribution of Xi given $Xi + Xj = r$, real life situations and applications.

3. Truncated Distributions: **(05L)**

Concept of truncated distribution, truncation to the right, left and on both sides. Binomial distribution left truncated at $X = 0$ (value zero is discarded), its p.m.f., mean and variance. Poisson distribution left truncated at $X = 0$ (value zero is discarded), its p.m.f., mean and variance. Real life situations and applications.

4 Time Series: **(14L)**

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

4.2 Exploratory data analysis: Time series plot to (i) check any trend and seasonality in the time series (ii) identify the nature of trend .

4.3 Methods of trend estimation and smoothing: (i) moving average, (ii) linear, parabolic, exponential, Parato curve fitting by least squares principle (iii) exponential smoothing.

4.4 Choosing parameters for smoothing and forecasting.

4.5 Forecasting based on exponential smoothing.

4.6 Measurement of seasonal variations: i) simple average method, ii) ratio to moving average method, iii) ratio to trend where linear trend is calculated by method of least squares.(Numerical examples with heavy computations are to be asked preferably in practicals).

4.7 Fitting of autoregressive model $AR(p)$, where $p = 1,2$.

4.8 Case studies of real life Time Series: Price index series, share price index series, economic time series: temperature and rainfall time series, wind speed time series, pollution levels.

****

SEMESTER – III

PAPER – II

ST 232 : CONTINUOUS PROBABILITY DISTRIBUTIONS

1. Continuous Univariate Distributions: (10L)

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 random variable.

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, mean deviation about mean.

1.3 Moment generating function (MGF): Definition, properties.
Cumulant generating function (CGF): Definition.

1.4 Mode, partition values : quartiles(Q_1, Q_2, Q_3), deciles, percentiles.

1.5 Probability distribution of function of r. v. : $Y = g(X)$ using 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 $g(X)$.

2. Continuous Bivariate Distributions: (09 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 random variables (events in terms of regions bounded by regular curves, circles, straight lines). Marginal and conditional distributions.

2.2 Expectation of r.v.(X,Y), expectation of function of r.v. $E[g(X,Y)]$, joint moments, $Cov(X,Y), Corr\ (X,Y)$, conditional mean, conditional variance, $E[E(X|Y=y)] = E(X)$ & $E[E(Y/X=x) = E(Y)$, regression as a conditional expectation.
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. $E(aX + bY + c), Var\ (aX + bY + c)$ (statement only proof not expected).

2.3 Independence of random variables $X\ and\ Y$ and also its extension to k random variables.

2.4 Moment generating function (MGF):$M_{X,Y}\ (t_1, t_2)$, properties, MGF of marginal distribution of random variables(r.v.s.), properties
i) $M_{X,Y}\ (t_1, t_2) = M_X\ (t_1, 0)\ M_Y\ (0, t_2)$ if X and Y are independent r.v.s.,

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

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

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

3. Standard Univariate Continuous Distributions:

3.1 Uniform or Rectangular Distribution: **(03 L)**

Probability density function (p.d.f.)

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

Notation : $X \sim U[a,b]$. p. d. f., sketch of p. d. f., c. d. f., mean, variance, symmetry, MGF.

Distributions 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 $X + Y, X - Y, XY$ and X/Y are not expected.)

3.2 Normal Distribution: **(10 L)**

Probability density function (p. d. f.)

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

Notation: $X \sim N(\mu\ \sigma^2)$.

p. d. f. curve, identification of scale and location parameters, nature of probability curve, mean, variance, MGF, CGF, central moments, cumulants, skewness, kurtosis, mode, quartiles(Q_1,Q_2,Q_3), points of inflexion of probability curve, mean deviation, additive property, probability distribution of : i) $\frac{X-\mu}{\sigma}$,standard normal variable (S.N.V.), ii) $aX + b$, iii) $aX + bY + c$, 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., computations of normal probabilities using normal probability integral tables. Central limit theorem (CLT) for $i.i.d.$ r.v.s. with finite positive variance(statement only), its illustration for Poisson and Binomial distributions.(Box-Muller transformation and normal probability plot to be covered in practicals)

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

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

Notation : $X \sim Exp(\alpha)$.

Nature of density curve, interpretation of α as a interarrival rate of custermers joining the queue and $\frac{1}{\alpha}$ as mean, mean, variance, MGF, CGF, skewness, kurtosis, c.d.f., graph of c.d.f., lack of memory property, quartiles(Q_1, Q_2, Q_3), mean deviation about mean, distribution of sum of two i.i.d exponential random variables. Distribution of $min(X,Y)$ and $max(X,Y)$ with X, Y $i.i.d.$ exponential random variables.

****

SEMESTER – III

PAPER - III

ST-233: PRACTICALS

Pre-requisites: Knowledge of the topics in theory.

Objectives:

1. To fit various discrete and continuous probability distributions and to study various real life situations.(Using calculators and $Ms - EXCEL$)
2. To identify the appropriate probability model that can be used.
3. To use forecasting and data analysis techniques in case of univariate data sets.

Preparation by Internal Examiner for Section I (Online examination):

1. Keep at least 15 **computers** with latest configuration ready with battery backup and necessary software, printers, scientific calculators, necessary statistical tables, normal probability paper at the examination laboratory.
2. Appropriate data set for time series, linear, quadratic, exponential trend fitting, exponential smoothing be entered in spreadsheet.
3. Any other type of data required as per practical slip 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 $Ms - EXCEL$ to operate on the data set which is already fed.
2. The question on section I are compulsory and there is no internal option.
3. The slips made available for Section I shall be allotted to the candiadtes at random so that the total marks of all asked commands will be exactly 07.

Sr. No.	Title of the Practical	No. of Practicals
1	Fitting of negative binomial distribution and computation of expected frequencies.	1
2	Fitting of normal distribution and computation of expected frequencies. Use of normal probability paper to check normality of raw data.	1
3	Applications of negative binomial and multinomial distributions.	1
4	Applications of normal distributions.	1
5	Model sampling from exponential distribution using distribution function, Model sampling from normal distribution using (i) distribution function, , (ii) Box-Muller transformation.	1
6	Time series : Estimation and forecasting of trend by exponential smoothing, moving averages, plotting of residuals. Fitting of $AR\ (1)$ model	1
7	Estimation of seasonal indices by (i)ratio to trend (ii)ratio to moving range	1
	Practicals based on MS- EXCEL	
8	Finding negative binomial probabilities. Fitting of negative binomial distribution and bar diagram of p.m.f. using $Ms-EXCEL$.	1
9	Finding normal probabilities .Fitting of normal distribution and tracing of normal probability curve using $Ms-EXCEL$.	1
10	Fitting of linear, quadratic , exponential trends to time series data. Finding the best fit using $R^{2.}$ Moving averages . Exponential smoothing using $Ms-EXCEL$.	2
11	**Project**: Project based on analysis of data collected by students in groups of maximum **6** students. (Project is equivalent to three practical's)	3
	Study tour report (if any)	-

****

SEMESTER – IV
PAPER – I
ST – 241: TESTS OF SIGNIFICANCE AND STATISTICAL METHODS.

1. Tests of Significance: (14L)

1.1 Random sample from a distribution as $i.i.d.$ r.vs. $X_1, X_2, X_3, \ldots\ldots, X_n$.

1.2 **Statistic and Parameter**. Sampling distribution of a statistic, standard error of a statistic with illustrations. **Statistical Inference**: Introduction to problem of Estimation and testing of hypothesis. Estimator and estimate. Unbiased estimator (definition and simple illustrations only). Point and interval estimation. Statistical hypothesis, null and alternative hypothesis, simple and composite hypothesis, one sided and two sided alternative hypothesis, critical region, $type - I$ and $type - II$ error, level of significance, $p - value$. Two sided confidence interval. Tests of hypotheses using i) critical region approach, $ii)$ $p - value$ approach and iii) confidence interval approach.

1.3 Tests for population means (large sample / approximate tests):

i) $H_0: \mu = \mu_0$ against $H_1: \mu \neq \mu_0$, $H_1: \mu > \mu_0$, $H_1: \mu < \mu_0$. (variance known)

ii) $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$, $H_1: \mu_1 > \mu_2$, $H_1: \mu_1 < \mu_2$. (variances known)

iii) Construction of two sided confidence interval for μ and $\mu_1 - \mu_2$

1.4 Tests for population proportions:

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

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

iii) Construction of two sided confidence interval for $P \; and \; P_1 - P_2$.

2. Multiple Linear Regression Model: (08L)

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

i) $0 \leq R_{Y.X_1X_2} \leq 1$ ii) $R_{Y.X_1X_2} \geq \min\{r_{yx_1}, r_{yx_2}\}$.

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

2.3 Partial correlation coefficient: Definition and derivation of partial correlation coefficient $r_{yx_1.x_2}$ and $r_{yx_2.x_1}$ Property of partial correlation coefficient ($-1 \leq r_{yx_1.x_2}, r_{yx_2.x_1} \leq 1$).(Statement only)

2.4 Notion of multiple linear regression. Yule's notation (trivariate case) (statement only).
Fitting 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, solution of normal equations. Definition and interpretation of partial regression coefficients β_1 and β_2. (relations between partial regression coefficients and multiple correlations on are not expected).

Residual: Definition, order, derivation of variance, properties. Finding multiple and partial correlation coefficients if (X_1, X_2, X_3) ~> MD(n, P_1, P_2 ,P_3)

3. Demography: (08L)

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

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

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

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

.(Numerical examples with heavy computations are to be asked preferably in practicals).

3.5 Interpretations of different rates, uses and applications.

3.6 Trends in vital rates as revealed in the latest census.

4. Queuing Model: (06L)

Introduction to quening model. as an application of exponential distribution, Poisson distribution and geometric distribution. Kendall's notation $M/M/1:FIFO/\infty/\infty$

Inter arrival rate (λ), service rate (μ), traffic intensity ($=\frac{\lambda}{\mu} < 1$),queue disciplines.

Probability distribution of number of customers in queue, average queue length, average waiting time in: i) queue, ii) system.(without derivations) statement of Little's formula / relations.

****

SEMESTER-IV

PAPER-II

ST-242: SAMPLING DISTRIBUTIONS AND EXACT TESTS.

1. **Gamma Distribution:** **(04 L)**

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

$= 0$, otherwise.

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

Nature of probability curve, special cases: i) $\alpha = 1$, ii) $\lambda = 1$,MGF, CGF, moments, cumulants, skewness, kurtosis, mode, additive property. Distribution of sum of n i.i.d. exponential variables. Relation between distribution function of Poisson and Gamma variates.

2. Chi-square Distribution: **(11 L)**

Definition of chisquare r.v. as a sum of squares of i.i.d. standard normal variables. Derivation of the p.d.f. of Chi-square variable with n degrees of freedom (d.f.) using MGF.

Notation: $X \sim \chi_n^2$

Mean, variance, MGF, CGF, central moments skewness, kurtosis, mode, additive property. Use of chi-square tables for calculations of probabilities. Normal approximation: $\frac{\chi_n^2 - n}{\sqrt{2n}}$ (statement only) Distribution of $\bar{X}$ and $\frac{nS^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=i}^{n}(X_i - \bar{X})^2$for a random sample from a normal distribution using orthogonal transformation, independence of $\bar{X}$ and S^2.

3. Student's t −distribution: **(05 L)**

Definition of t r.v. with n d.f. in the form of $t = \frac{U}{\sqrt{\frac{V}{n}}}$, where $U \sim N(0,1)$ and V is chi-square with n d.f., where U & V are independent random variables.

Notation: $t \sim t_n$

Derivation of the p.d.f of t distribution, nature of probability curve, mean, variance, moments, mode. Use of t-tables for calculations of probabilities, statement of normal approximation.

4. Snedecore's F −distribution: **(06 L)**

Definition of F r.v. with n_1 and n_2 d.f. as $F_{n_1,n_2} = \frac{X_1/n_1}{X_2/n_2}$ where X_1 & X_2 are independent chi-square variables with n_1 and n_2 d.f.

Notation: $F \sim F_{n_1,n_2}$

Derivation of the p.d.f, nature of probability curve, mean, variance, moments, mode.

Distribution of $\frac{1}{F_{n_1,n_2}}$,use of F −tables for calculation of probabilities.

Interrelationship between Chi-square, t and F distributions.

5. Test of Hypothesis: (10 L)

5.1 Tests based on chi-square distribution:

a) Test for independence of two attributes arranged in 2×2 contingency table(with Yate's correction) (to be covered in practical only)

b) Test for independence of two attributes arranged in $r \times s$ contingency table, Mc Nemar's test (to be covered in practical only)

c) Test for goodness of fit. (to be covered in practical only)

d) Test for variance (Ho: $\sigma^2 = \sigma_o^2$)against one-sided and two-sided alternatives i) for known mean , ii) for unknown mean.

5.2 Tests based on t −distribution:

a) Tests for population means:

(i) Single sample with unknown variance and two sample for unknown equal variances tests for one-sided and two-sided alternatives.

(ii)$100(1-\alpha)\%$ two sided confidence interval for population mean and difference of means of two independent normal populations.

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

5.3 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 and ii) means are unknown. Take F= $\frac{S_1^2}{S_2^2}$.

****

SEMESTER IV

PAPER III

ST-243: PRACTICALS

Pre-requisites : Knowledge of the topics in theory.

Objectives:

1. To conduct various tests of significance like averages, population proportions, independence of attributes, variance etc. included in theory (using calculators, *R* software).
2. To compute probabilities of discrete and continuous probability distributions using R software.
3. To use *R* software for finding basic summary statistics.

Preparation by Internal Examiner for Section I (Online examination):

1. Keep at least 15 **computers** with latest configuration ready with battery backup and necessary software, printers, scientific calculators, necessary statistical tables at the examination laboratory.

Instructions to Examiners:

1. The question on section I are compulsory and there is no internal option.
2. The commands of the nature attached in specimen are to be asked, so that the total marks of all asked commands will be exactly 7.

Sr. No.	Title of the Practical	No. of Practicals
1	Computations of GRR and NRR	1
2	Test for proportions and construction of confidence interval for $H_0: P = P_0 \; and \; H_0: P_1 = P_2$	1
3	Test for means and construction of confidence interval i) $H_0: \mu = \mu_0$, σ^2 known and σ^2 unknown ii) (ii) $H_0: \mu_1 = \mu_2$, σ_1^2 and σ_2^2 both known $iii) H_0: \mu_1 = \mu_2$, $\sigma_1^2 = \sigma_2^2 = \sigma^2$ unknown (iv) $H_0: \mu_1 = \mu_2$ (paired t test)	1
4	Tests based on χ^2 distribution (i)Goodness of fit i) Independence of attributes (2 × 2, r × s contingency table) ii) Mc Nemar's test iii) $H_0: \sigma^2 = \sigma_0^2$, μ unknown, confidence interval for σ^2	2
5	Tests based on F distribution $H_0: \sigma_1^2 = \sigma_2^2$ for i) means known ii) means unknown	1
	Practicals based on R software	
6	Use of basic R software commands c( ), scan( ), rep( ), seq( ), min, max, sort, extract, data.frame, matrix, accessing resident data sets etc.	1
7	Finding summary statistics using summary ( ) and fivenum( ). Calculate arithmetic mean (AM), geometric mean (GM), harmonic mean (HM), median, mode, quantiles, range, quartile deviation (QD), variance, coefficient of variation (CV) using R software.	1
8	Computation of probabilities of negative binomial, multinomial, normal, exponential, gamma, t, χ^2, F using R software	1
9	Tests for proportions, means, χ^2 distribution, F distribution using R software. Fitting of trivariate regression plane using R software	1
10	**Project**: Project based on analysis of data collected by students in groups of maximum **6** students. (Project is equivalent to three practical's)	3
	Study tour report (if any)	-

****

Books Recommended:

1. Brockwell P.J.and Davis R.A. (2003), Introduction to Time Series and Forecasting (Second Edition), Springer Texts in Statistics.
2. Chatfield C. (2001), The Analysis of Time Series An Introduction, Chapman and Hall / CRC, Texts in Statistical Science .
3. Goon A. M., Gupta, M. K. and Dasgupta, B. (1986), Fundamentals of Statistics, Vol. 2, World Press, Kolkata.
4. Gupta, S. C. and Kapoor, V. K. (2002), Fundamentals of Mathematical Statistics, (Eleventh Edition), Sultan Chand and Sons, 23, Daryaganj, New Delhi , 110002 .
5. Gupta, S. C. and Kapoor V. K. (2007), Fundamentals of Applied Statistics (Fourth Edition), Sultan Chand and Sons, New Delhi.
6. Gupta, S. P. (2002), Statistical Methods (Thirty First Edition), Sultan Chand and Sons, 23, Daryaganj, New Delhi 110002.
7. Hogg, R. V. and Craig, A. T. , Mckean J. W. (2012), Introduction to Mathematical Statistics (Tenth Impression), Pearson Prentice Hall.
8. Kulkarni, M. B., Ghatpande, S. B. and Gore, S. D. (1999), Common Statistical Tests, Satyajeet Prakashan, Pune 411029
9. Medhi, J., Statistical Methods, Wiley Eastern Ltd., 4835/24, Ansari Road, Daryaganj, New Delhi – 110002.
10. Meyer, P. L., Introductory Probability and Statistical Applications, Oxford and IBH Publishing Co. New Delhi.
11. Mood, A. M., Graybill F. A. and Bose, F. A. (1974), Introduction to Theory of Statistics (Third Edition, Chapters II, IV, V, VI), McGraw - Hill Series G A 276
12. Mukhopadhya Parimal (1999), Applied Statistics, New Central Book Agency, Pvt. Ltd. Kolkata
13. Purohit S. G., Gore S. D. and Deshmukh S. R. (2008), Statistics using R, Narosa Publishing House, New Delhi.
14. Ross, S. (2003), A first course in probability (Sixth Edition), Pearson Education publishers , Delhi, India.
15. Walpole R. E., Myers R. H. and Myers S. L. (1985), Probability and Statistics for Engineers and Scientists (Third Edition, Chapters 4, 5, 6, 8, 10), Macmillan Publishing Co. Inc. 866, Third Avenue, New York 10022.
16. Weiss N., Introductory Statistics, Pearson education publishers.