ST. JOSEPH'S COLLEGE (AUTONOMOUS)

BENGALURU-27

Re-accredited with **'A++' GRADE with 3.79/4 CGPA** by
NAAC Recognized by UGC as College of Excellence

DEPARTMENT OF STATISTICS

Syllabus for Bachelor of Science (Basic/Hons.)

With Statistics as one of the majors with practical with another subject
as another major in 3rd year

Under National Education Policy

For Batch 2021 (Onwards)

Name of the Degree Program: B.Sc.
Discipline Core: Statistics
Total Credits for the Program: 176 (till 8 semesters)
Starting year of implementation: 2021-22

Program Outcomes

By the end of the program the students will be able to:

1. Acquire fundamental/systematic or coherent understanding of the academic field of Statistics and its different learning areas and applications.
2. Develop and demonstrate an ability to understand major concepts in various disciplines of Statistics.
3. Demonstrate the ability to use skills in Statistics and different practicing areas for formulating and tackling Statistics related problems and identifying and applying appropriate principles and methodologies to solve a wide range of problems associated with Statistics.
4. Understand procedural knowledge that creates different types of professionals related to subject area of Statistics, including professionals engaged in government/public service and private sectors.
5. Plan and execute Statistical experiments or investigations, analyze and interpret data/information collected using appropriate methods, including the use of appropriate statistical software including programming languages, and report accurately the findings of the experiment/investigations.
6. Have a knowledge regarding use of data analytics tools like Excel and R-programming.
7. Developed ability to critically assess a standard report having graphics, probability statements.
8. Analyze, interpret the data and hence help policy makers to take a proper decision.
9. Recognize the importance of statistical modelling and computing, and the role of approximation and mathematical approaches to analyze the real problems using various statistical tools.
10. Demonstrate relevant generic skills and global competencies such as
    i. Problem-solving skills that are required to solve different types of Statistics related problems with well-defined solutions, and tackle open-ended problems, that belong to the disciplinary-area boundaries;
    ii. Investigative skills, including skills of independent thinking of Statistics-related issues and problems;
    iii. Communication skills involving the ability to listen carefully, to read texts and reference material analytically and to present information in a concise manner to different groups/audiences of technical or popular nature;
    iv. Analytical skills involving paying attention to details and ability to construct logical Arguments using correct technical language related to Statistics and ability to translate them with popular language when needed;
    v. ICT skills;
    vi. Personal skills such as the ability to work both independently and in a group.

11. Undertake research projects by using research skills- preparation of questionnaire, conducting national sample survey, research projects using sample survey, sampling techniques.
12. Understand and apply principles of least squares to fit a model to the given data, study the association between the variables, applications of Probability Theory and Probability Distributions.

Assessment

Weightage for assessments

Type of Course	CA Marks	SE Marks	Total Marks
Theory	40	60	100
Practical	20	30	50
Open elective	40	60	100

SE Marks

Type of Course	Max marks for ESE	Reduced Marks	Test Duration
Theory	100	60	3 Hrs
Practical	30	30	3 Hrs
Open elective	60	60	2Hrs

CA Marks

Title	Maximum Marks	Reduced Marks
Activity 1	10	10
Mid semester test	50	20
Activity 2	10	10
Activity 3	Optional (For improvement)	
Total		40

Model Program Structures for the Under-Graduate Programs in Universities and Colleges in Karnataka

Bachelor of Arts (Basic/ Hons.)/ Bachelor of Science (Basic/ Hons.) etc. with Statistics as Major with practical and another subject as major

Sem.	Discipline Core (DSC) (Credits) (L+T+P)	Discipline Elective (DSE) /Open Elective (OE) (Credits) (L+T+P)	Ability Enhancement Compulsory Courses (AECC), Languages (Credits)(L+T+P)		Skill Enhancement Courses (SEC)		Total Credits
					Skill based (Credits) (L+T+P)	Value based (Credits) (L+T+P)	
I	**Descriptive Statistics** (4+2) Discipline B1(4+2)	OE-1 (3)	L1-1 (3), L2-1 (3)(3+1+0 each)		SEC-1: Digital Fluency (2) (1+0+2)		23
II	**Probability and Distributions** (4+2) Discipline B2(4+2)	OE-2 (3)	L1-2(3), L2-2 (3) (3+1+0 each)	Environmental Studies (2)		Health & Wellness/ Social & Emotional Learning (2) (1+0+2)	25
Exit option with Certificate (48 credits)							
III	**Calculus and Probability Distributions** (4+2) Discipline B3(4+2)	OE-3 (3)	L1-3 (3), L2-3(3) (3+1+0 each)		SEC-2: Artificial Intelligence (2)(1+0+2)		23
IV	**Statistical Inference-I** (4+2) Discipline B4(4+2)	OE-4 (3)	L1-4 (3), L2-4(3) (3+1+0 each)	Constitution of India (2)		Sports/NCC/NSS etc. (2) (1+0+2)	25
Exit option with Diploma (96 credits)							
V	**Matrix Algebra and Regression Analysis** (3+2) **Analysis of variance and design of experiments** (3+2) Discipline B5(3+2)	DS-B Elective 1 (3)			SEC-3: Cyber Security (2) (1+0+2)	Ethics & Self Awareness (2) (1+0+2)?	20

VI	**Statistical Inference-II**(3+2) DisciplineB6(3+2) DisciplineB7(3+2)	DS-A Elective 1 (3)			SEC-4: Professional/ Societal Communication (2)		20
Exit option with Bachelor of Arts, B.A. / Bachelor of Science, B. Sc. Basic Degree (136 credits)							
Choose any one Discipline as Major							
VII	**Sample Surveys and Statistics for National Development** (3+2) **Real Analysis** (3+2) **Probability Theory** (4)	DS-A/B Elective 2(3) Res. Methodology (3)					20
VIII	**Linear Algebra** (4) **Linear models and Design of Experiments** (4)	DS-A/B Elective 3(3) DS-A/B Elective 4(3) Research Project (6)*					20
Award of Bachelor of Arts Honours, B.A. (Hons.)/ Bachelor of Science Honours, B.Sc. (Hons) degree in a discipline etc. (176 credits)							
IX	**Multivariate Analysis (3)+ Practical (2)** **Decision Theory and Bayesian Inference(4)** **Distribution Theory (3) +Practical (2)**	Statistics E-5(3:0:0) Research Methodology (3)					20
X	**Stochastic Processes (4)** **Time Series Analysis (3)** **Machine Learning (4)**	Statistics E-6 (3:0:0) Research Project (6)					20
Award of Master of Science Degree in Statistics							

Summary of Discipline Specific Courses (DSC)			
Semester	**Course Code**	**Title of the Paper**	**Credits**
I	ST 121	Descriptive Statistics	4
	ST 1P1	Practicals based on ST 121	2
II	ST 221	Probability and Distributions	4
	ST 2P1	Practicals based on ST 221	2
III	ST 321	Calculus and Probability Distributions	4
	ST 3P1	Practicals based on ST 321	2
IV	ST 421	Statistical Inference-I	4
	ST 4P1	Practicals based on ST 421	2
V	ST 5121	Matrix Algebra and Regression Analysis	3
	ST 5P1	Practicals based on ST 5121	2
	ST 5221	Analysis of variance and design of experiments	3
	ST 5P2	Practicals based on ST 5221	2
VI	ST 6121	Statistical Inference-II	3
	ST 6P1	Practicals based on ST 6121	2
VII	ST 7121	Sample Surveys and Statistics for National Development	3
	ST 7P1	Practicals based on ST 7121	2
	ST 7221	Real Analysis	3
	ST 7P2	Practicals based on ST 7221	2
	ST 7321	Probability Theory	4
VIII	ST 8121	Linear Algebra	4
	ST 8221	Linear models and Design of Experiments	4

Semester	Course Code	Title of the Paper	Credits
IX	ST 9121	Multivariate Analysis	3
	ST 9P1	Practical based on ST 9121	2
	ST 9221	Decision Theory and Bayesian Inference	4
	ST 9321	Distribution Theory	3
	ST 9P2	Practical based on ST 9321	2
X	ST 0121	Stochastic Processes	4
	ST 0221	Time Series Analysis	3
	ST 0321	Machine Learning	2

List of Discipline Specific Electives (DSE)

- Actuarial Statistics
- Advanced Statistical Inference
- Analysis of Categorical Data
- Analysis of Clinical Trials
- Artificial Intelligence with R
- Bayesian Inference
- Bio-Statistics
- Computational Statistics
- Data Analytics with R/Python
- Data Science: Multivariate Techniques with R /Python
- Data Science with R/Python
- Demography
- Extreme value Theory
- Econometrics
- Financial Statistics
- Multivariate Techniques
- Nonparametric and Semiparametric Methods
- Operations Research
- Project Work
- Reliability Analysis
- Reliability and Statistical Quality Control
- Statistical Learning and Data Mining with R/Python
- Statistical Quality Control
- Stochastic Models in Finance
- Survival Analysis
- Time Series Analysis
- Sampling Theory and Applications

List of Open Elective (OE) for first two semesters

- Statistical Methods
- Business Statistics
- Applied statistics
- Biostatistics

Curriculum Structure for the Undergraduate Degree Program

B.Sc.

Total Credits for theProgram:176

Starting year of implementation:2021-22

Name of the Degree Program: B.Sc.

Discipline/Subject: Statistics (Major)

Program Articulation Matrix

This matrix lists only the core courses. Core courses are essential to earn the degree in that discipline/subject. They include courses such as theory, laboratory, project, internships etc. Elective courses may be listed separately

Semester	Title /Name of the course	Program outcomes that the course addresses (not more than 3 per course)	Pre-requisite course(s)	Pedagogy[##]	Assessment[$]
1	**Descriptive Statistics**	PO1,PO2,PO8	Mathematics of 12th level	1. The course is taught using traditional chalk and talk method using problem solving through examples and exercises. 2. Students are encouraged to use resources available on open sources.	The assessment is done using continuous assessment through written test, open book examination, viva-voce, seminars, and group discussions.
1	**Practical**	PO5, PO6	Mathematics of 12th level	The course is taught using Excel software and/or manually to carry out descriptive statistical analysis.	Assessment of learning through experiments
2	**Probability and Distributions**	PO7,PO9,PO10	Mathematics of 12th level	1. The course is taught using traditional chalk and talk method using problem solving through examples and exercises. 2. Students are encouraged to use resources available on open sources	The assessment is done using continuous assessment through written test, open book examination, viva-voce, seminars, and group discussions.
2	**Practical**	PO5, PO6	Mathematics of 12th level	The course is taught using R programming software and/or manually to carry out descriptive statistical analysis	Assessment of learning through experiments

[##] Pedagogy for student engagement is predominantly lectures. However, other pedagogies enhancing better student engagement to be recommended for each course. The list includes active learning/course projects/ problem or project-based learning/ case studies/self-study like seminar, term paper or MOOC

[$] Every course needs to include assessment for higher order thinking skills (Applying/ analyzing/ Evaluating/ Creating). However, this column may contain alternate assessment methods that help formative assessment (i.e., assessment for learning).

Course Pre-requisite(s): II PUC with Mathematics

Course Outcomes (COs)

At the end of the course the student should be able to:

1. Acquire knowledge of introductory statistics, its scope and importance in various areas such as Medical, Engineering, Agricultural and Social Sciences etc.
2. Get knowledge of various types of data, their organization and evaluation of summary measures such as measures of central tendency and dispersion etc.
3. Perceive the knowledge of correlation, regression analysis, regression diagnostics, partial and multiple correlations.
4. Learn different of types of data reflecting independence and association between two or more attributes.
5. Develop ability to critically assess a standard report having graphics, probability statements.
6. Conceptualize the probabilities of events including frequentist and axiomatic approach. Simultaneously, they will learn the notion of conditional probability including the concept of Bayes' Theorem,
7. Get knowledge related to concept of discrete and continuous random variables and their probability distributions including expectation and moments,
8. Learn knowledge of important discrete and continuous distributions such as Binomial, Poisson, Normal distributions.
9. Acquire knowledge on R-programming in the descriptive statistics and probability models.

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs 1 – 12)

Course Outcomes (COs) / Program Outcomes (POs)	1	2	3	4	5	6	7	8	9	10	11	12
1. Knowledge of introductory statistics, its scope and importance in various areas such as Medical, Engineering, Agricultural and Social Sciences etc.	X	X			X	X						
2.Knowledge of various types of data, their organization and evaluation of summary measures such as measures of central tendency and dispersion etc.			X	X	X	X				X	X	
3. Knowledge of correlation, regression analysis, regression diagnostics, partial and multiple correlations.				X	X	X		X		X	X	
4.Knowledge of types of data reflecting independence and association between two or more attributes				X	X	X				X		X
5. Develop ability to critically assess a standard report having graphics, probability statements.					X	X	X		X			
6. Knowledge to conceptualize the probabilities of events including frequentist and axiomatic approach. Simultaneously, they will learn the notion of conditional probability including the concept of Bayes' Theorem.					X	X			X	X		
7. Knowledge related to concept of discrete and continuous random variables and their probability distributions including expectation and moments.					X	X			X	X		
8. Knowledge of important discrete and continuous distributions such as Binomial, Poisson, Normal, distributions.					X	X			X	X		
9. Knowledge on R-programming in the descriptive statistics and probability models.					X	X			X	X		

Course Articulation Matrix relates course outcomes of course with the corresponding program outcomes whose attainment is attempted in this course. 'X' in the intersection cell indicates that particular course outcome addresses that particular program outcome.

Bachelor of Science (Basic/Hons.) in Statistics

First Semester

ST – 121: Descriptive Statistics

Semester	**I**
Paper Code	**ST-121**
Paper Title	**Descriptive Statistics**
Number of teaching hours per week	04
Total number of teaching hours per semester	56
Number of credits	04
Formative Assessment Marks	40
Summative Assessment Marks	60

Unit – 1: Introduction to Statistics 13 Hours

Statistics: Definition and scope. Concepts of statistical population and sample. Data: quantitative and qualitative, cross sectional and time-series, discrete and continuous. Scales of measurement: nominal, ordinal, interval and ratio. Samplin theory concepts: (SRS, Stratified, Systematic and Cluster sampling methods (Definitions only). Presentation of data: tabular and graphical. Frequency distributions, cumulative frequency distributions and their graphical representations. Stem and leaf displays.

Unit – 2: Univariate Data Analysis 18 Hours

Measures of Central Tendency: Mean, weighted mean, trimmed mean, Median, Mode, Geometric and harmonic means, properties, merits and limitations, relation between these measures. Measures of Dispersion: Range, Quartile deviation, Mean deviation, Standard deviation and their relative measures. Moments, Skewness and Kurtosis. Quantiles and measures based on them. Box Plot. Outliers. Chebyshev's inequality, normal data sets.

Unit – 3: Bivariate Data Analysis 15 Hours

Bivariate Data, Scatter plot, Correlation, Karl Pearson's correlation coefficient, Rank correlation – Spearman's and Kendall's measures. Concept of errors, Principle of least squares, fitting of polynomial and exponential curves. Simple linear regression and its properties. Fitting of linear regression line and coefficient of determination.

Unit –4: Multivariate Data Analysis 10 Hours

Analysis of Categorical Data: Contingency table, independence and association of attributes, measures of association - odds ratio, Pearson's and Yule's measure, Multivariate data sets and its visualization, illustration of mean vector and dispersion matrix, Multiple linear regression, multiple and partial correlation coefficients (Only for 3 variables). Residual error variance.

References

1. Agresti, A. (2010): Analysis of Ordinal Categorical Data, 2nd Edition,Wiley.
2. Anderson T.W. and Jeremy D. Finn (1996). The New Statistical Analysis of Data, Springer
3. Freedman,D.,Pisani,R.andPurves,R.(2014),Statistics,4thEdition,W.W.Norton& Company.
4. Gupta, S.C. (2018), Fundamental of Statistics, Himalaya Publishing House, 7thEdition.
5. Gupta S.C. and V.K. Kapoor (2020), Fundamental of Mathematical Statistics, Sultan Chand and Co. 12th Edition.
6. Hogg, R. V. McKean J. W. and Craig, A. T. (2012), Introduction to Mathematical Statistics, Pearson 7th Edition.
7. Joao Mendes Moreira, Andre C P L F de Carvalho, Tomas Horvath (2018), General Introduction to Data Analytics, Wiley.
8. Johnson, R.A. and Bhattacharyya, G.K. (2006), Statistics: Principles and methods. 5th Edition, John Wiley & Sons, New York.
9. Medhi, J. (2005), Statistical Methods, New Age International.
10. Ross, S.M. (2014), Introduction to Probability and Statistics for Engineers and Scientists, 5th Edition, Academic Press.
11. Tukey, J.W. (1977), Exploratory Data Analysis, Addison-Wesley Publishing Co.

Pedagogy

1. The course is taught using traditional chalk and talk method using problem solving through examples and exercises.
2. Students are encouraged to use resources available on open sources.

MODEL BLUEPRINT		
Code number:		**ST – 121**
Title of the paper:		**Introduction to Statistics and Probability**
Unit no	**No of hrs**	**Total marks**
1	13	30
2	18	42
3	15	35
4	10	23
Total	56	130
Maximum marks Descriptive (Excluding Bonus Questions)		80
Maximum marks Objective		20
Maximum Marks (Excluding Bonus Questions)		100

Practical I
ST – 1P1– Practical on ST – 121

Number of teaching hours per week	04
Total number of teaching hours per semester	52
Number of credits	02
Formative Assessment Marks	20
Summative Assessment Marks	30

Computing all the practical manually and using Excel

1. Presentation of data by frequency tables, diagrams and graphs, stem and leaf, partition values.
2. Arithmetic Mean (AM), geometric mean, harmonic mean, weighted AM, trimmed mean, corrected mean.
3. Mode, median, partition values.
4. Absolute and relative measures of dispersion, Box plots.
5. Problems on moments, skewness and kurtosis.
6. Fitting of curves by least squares method.
7. Product moment correlation coefficient and rank correlation.
8. Regression of two variables.
9. Multivariate Descriptive statistics, mean vector, dispersion matrix, correlation matrix, Partial and Multiple correlation.
10. Problems on Association of attributes

Bachelor of Science (Basic/Hons.) in Statistics

Second Semester

ST 221: Probability and Distributions

Semester	**II**
Paper Code	**ST-221**
Paper Title	**Probability and Distributions**
Number of teaching hours per week	04
Total number of teaching hours per semester	56
Number of credits	04
Formative Assessment Marks:	40
Summative Assessment Marks:	60

Unit –1: Probability 15 Hours

Random experiment, sample space and events, algebra of events. Definitions of Probability- Classical, statistical, subjective and axiomatic approaches – illustrations and applications,

Addition Theorem, Conditional probability, independence of events and multiplication Theorem, Total probability rule, Bayes theorem and applications.

Unit –2: Random Variables and Mathematical Expectation 14 Hours

Definitions of discrete and continuous random variables, Distribution function, probability mass and density functions – properties and illustrations, Expectation of a random variable and Properties of expectation, Moments and moment generating function – properties and uses.

Unit –3: Standard Distributions 16 Hours

Bernoulli, Binomial, Poisson, Geometric distributions– mean, variance, moments and m. g. f. recursive relations for probabilities and moments of Binomial and Poisson distributions, Uniform, Exponential and Normal distribution: Mean and Variance, m. g. f. Memoryless property, Area property of normal distribution. Fitting of distributions and to assessing its goodness of fit (Graphically).

Unit –4: Data Analysis Using R 11 Hours

Introduction to R: Installation, command line environment, overview of capabilities, brief mention of open-source philosophy. R as a calculator: The four basic arithmetic operations. Use of parentheses nesting up to arbitrary level. The power operation. Evaluation of simple expressions. Quotient and remainder operations for integers. Standard functions, e.g., sin, cos, exp, log. The different types of numbers in R: Division by zero leading to Inf or -Inf. NaN. NA. No need to go into details. Variables. Creating a vector using c(), seq() and colon operator. How functions map over vectors. Functions to summarize a vector: sum, mean, sd, median etc. Extracting a subset from the vector (by index, by property). R as a graphing calculator: Introduction to plotting. Plot(), lines(), abline(). No details about the graphics parameters except colour and line width. Barplot, Pie chart and Histogram. Box plot. Scatter plot and simple linear regression using lm(y~x). Problems on discrete and continuous probability distributions

References

1. Dudewitz. E.J. and Mishra. S. N. (1998), Modern Mathematical Statistics. John Wiley.
2. Goon A.M., Gupta M.K., Das Gupta.B. (1991), Fundamentals of Statistics, Vol. I, World Press, Calcutta.
3. Gupta. S.C and V.K. Kapoor (2020), Fundamentals of Mathematical Statistics, Sultan Chand and Co, 12th Edition.
4. Hogg, R.V., Tanis, E.A. and Rao J.M. (2009), Probability and Statistical Inference, Seventh Edition, Pearson Education, New Delhi.
5. Mood, A.M., Graybill, F.A. and Boes, D.C. (2007), Introduction to the Theory of Statistics, 3rd Edition. (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
6. Ross, S. (2002), A First Course in Probability, Prentice Hall.
7. Sudha G. Purohit, Sharad D. Gore, Shailaja R Deshmukh, (2009), Statistics using R, Narosa Publishing House.
8. R for beginners by Emmanuel Paradis (freely available at https://cran.r-project.org/doc/contrib/Paradisrdebuts_en.pdf)

Pedagogy

1. The course is taught using traditional chalk and talk method using problem solving through examples and exercises.
2. Students are encouraged to use resources available on open sources.

MODEL BLUEPRINT		
Code number:		**ST – 221**
Title of the paper:		**Probability and Distribution**
Unit number	**Number of hrs**	**Total marks**
1	15	35
2	14	32
3	16	35
4	11	28
Total	56	130
Maximum marks Descriptive (Excluding Bonus questions)		80
Maximum marks Objective		20
Maximum marks (Excluding Bonus questions)		100

Practical II

ST – 2P1– Practical on ST – 221

Number of teaching hours per week	04
Total number of teaching hours per semester	52
Number of credits	02
Formative Assessment Marks	20
Summative Assessment Marks	30

Computing all the practical manually and using Excel/R

1. Two exercises on Descriptive statistics (Presentations, Summarizations, correlations, regression and Graphs using R)
2. Computing probability: using addition and multiplication theorems.
3. Conditional probability and Bayes' theorem.
4. Problems on pmf, expectation, variance, quantiles, skewness, kurtosis (Discrete Case).
5. Problems on pdf, expectation, variance, quantiles, skewness, kurtosis (Continuous case).
6. Problems on discrete probability distributions (Binomial, Poisson, Geometric)
7. Problems on Uniform, Exponential and Normal probability distributions
8. Computation of moments and Moment generating functions (Discrete and Continuous Case).
9. Fitting of distributions (Binomial, Poisson, Geometric, Uniform, Exponential, Normal)
10. Generation of random samples. (Binomial, Poisson, Geometric, Uniform, Exponential, Normal)

****

STOE: 1 Statistical Methods

Course Objectives

1. This is an open elective course for other than statistics students.
2. The students will learn the elements of descriptive statistics, probability, statistical methods such as tests of hypotheses, correlation and regression.

Course Pre-requisite: II PUC or equivalents with Mathematics

Course Outcomes

Upon the completion of this course students should be able to

1. CO1: Acquire knowledge of statistical methods.
2. CO2: Identify types of data and visualization, analysis and interpretation.
3. CO3: Know about elementary probability and probability models.
4. CO4: Employ suitable test procedures for given data set.

Pedagogy

The course is taught using traditional chalk and talk method using problem solving through examples and exercises. Students are encouraged to use resources available on open sources.

Unit 1: Introduction 10 Hours

Definition and scope of Statistics. Data: quantitative and qualitative, attributes, variables, scales of measurement - nominal, ordinal, interval and ratio. Presentation: tabular and graphic, including histogram and ogives. Concepts of statistical population and sample. Sampling from finite population - Simple random sampling, Stratified and systematic random sampling procedures (definitions and methods only). Concepts of sampling and non- sampling errors.

Unit 2: Univariate and Bivariate Data Analysis 10 Hours

Measures of Central Tendency: mathematical and positional. Measures of Dispersion: range, quartile deviation, mean deviation, standard deviation, coefficient of variation, moments, skewness and kurtosis.

Bivariate data, scatter diagram, Correlation, Karl-Pearson's correlation coefficient, Rank correlation. Simple linear regression, principle of least squares and fitting of polynomials

Unit 3: Probability and Distributions 12 Hours

Probability: Random experiment, trial, and sample space, events-mutually exclusive and exhaustive events. Classical, statistical and axiomatic definitions of probability, addition and multiplication theorems, Bayes theorem (only statements). Discrete and continuous random variables, probability mass and density functions, distribution functions, expectation of a random variable. Standard univariate distributions: Binomial, Poisson and Normal distributions (Elementary properties and applications only).

Unit 4: Sampling Distributions and Testing of Hypothesis 10 Hours

Distribution of sample mean from a normal population, Chi-square, t and F distributions (No derivations) and their applications.

Statistical Hypothesis - null and alternative hypothesis, simple and composite hypothesis.

Type- I and Type-II errors, level of significance, critical region, P-value and its interpretation.

Test for single mean, equality of two means, tests for Proportions, Chi-square goodness of fit, One-way ANOVA

References

1. Daniel, W. W. (2007 Biostatistics - A Foundation for Analysis in the Health Sciences, Wiley
2. T.W. Anderson and Jeremy D. Finn(1996). The New Statistical Analysis of Data, Springer.
3. Mukhopadhyay P(l 999). Applied Statistics, New Central book Agency, Calcutta.
4. Ross, S.M. (2014) Introduction to Probability and Statistics for Engineers and Scientists.
5. Cochran, W G (1984): Sampling Techniques, Wiley Eastern, New Delhi.

MODEL BLUEPRINT		
Code number:		**STOE – 1**
Title of the paper:		**Statistical methods**
Unit number	**Number of hrs.**	**Total marks**
1	10	20
2	10	20
3	12	25
4	10	20
Total	42	85
Maximum marks (Excluding bonus questions)		60

STOE – 2: Business Statistics

Course Objectives

1. Provide an introduction to basics of statistics within a financial context.
2. To enable students to use statistical techniques for analysis and interpretation of business data.

Course Pre-requisite: II PUC or equivalents

Course Outcomes

Upon the completion of this course students should be able to:

1. CO1: Frame and formulate management decision problems.
2. CO2: Understand the basic concepts underlying quantitative analysis.
3. CO3: Use sound judgment in the applications of quantitative methods to management decisions.

Pedagogy

1. The course is taught using traditional chalk and talk method using problem solving through examples and exercises.
2. Students are encouraged to use resources available on open sources.

Unit 1: Statistical Data and Descriptive Statistics 12 Hours

Introduction, Definition and Scope of Statistics.
Nature and Classification of data: univariate, bivariate and multivariate data; time-series and cross- sectional data.
Measures of Central Tendency: mathematical averages including arithmetic mean geometric mean and harmonic mean, properties and applications. Positional Averages Mode and Median (and other partition values including quartiles, deciles, and percentiles).
Measures of Variation: absolute and relative. Range, quartile deviation, mean deviation, standard deviation, and their coefficients, Properties of standard deviation/variance Skewness: Meaning, Measurement using Karl Pearson and Bowley's measures; Concept of Kurtosis.

Unit 2: Simple Correlation and Regression Analysis 10 Hours

Correlation Analysis: Meaning of Correlation: simple, multiple and partial; linear and non-linear, Correlation and Causation, Scatter diagram, Pearson's co-efficient of correlation; calculation and properties (Proof not required). Correlation and Probable error; Rank Correlation.
Regression Analysis: Principle of least squares and regression lines, Regression equations and estimation; Properties of regression coefficients; Relationship between Correlation and Regression coefficients; Standard Error of Estimate and its use in interpreting the results.

Unit 3: Index Numbers 10 Hours

Definition, Problems involved in the construction of index numbers, methods of constructing index numbers of prices and quantities , simple aggregate and price relatives method, weighted aggregate and weighted average of relatives method, important types of weighted index numbers: Laspeyre's, Paasche's, Bowley's, Marshall- Edgeworth, Fisher' s, method of obtaining price and quantity index numbers, tests consistency of index numbers, time reversal test and

factor reversal test for index numbers, Uses and limitations of index numbers. Consumer price index number: Problems involved in the construction of cost-of-living index number, advantages and disadvantages, Aggregative expenditure method and Family budget method for the construction of consumer price index numbers. Definition and measurement of Inflation rate - CPI and GNP Deflator.

Unit 4: Time Series Analysis 10 Hours

Introduction, definition and components of Time series, illustrations, Additive, Multiplicative and mixed models, analysis of time series, methods of studying time series: Secular trend, method of moving averages, least squares method - linear, quadratic, exponential trend fittings to the data. Seasonal variation - definition, illustrations, measurements, simple average method, ratio to moving average method, ratio of trend method, link relatives' method, Cyclical variation- definition, distinction from seasonal variation, Irregular variation- definition, illustrations.

References

1. Levin, Richard, David S. Rubin, Sanjay Rastogi, and HM Siddiqui. Statistics for Management. 7th ed., Pearson Education.
2. David M. Levine, Mark L. Berenson, Timothy C. Krehbiel, P. K. Viswanathan, Business Statistics: A First Course, Pearson Education.
3. Siegel Andrew F. Practical Business Statistics. McGraw Hill Education.
4. Gupta, S.P., and Archana Agarwal. Business Statistics, Sultan Chand and Sons, New Delhi.
5. Vohra N. D., Business Statistics, McGraw Hill Education.
6. Murray R Spiegel, Larry J. Stephens, Narinder Kumar. Statistics (Schaum's Outline Series), Mc-Graw Hill Education.
7. Gupta, S.C. Fundamentals of Statistics. Himalaya Publishing House.
8. Anderson, Sweeney, and Williams, Statistics for Students of Economics and Business, Cengage Learning

MODEL BLUEPRINT		
Code number:		**STOE – 2**
Title of the paper:		**Business Statistics**
Unit number	**Number of hrs**	**Total marks**
1	12	25
2	10	20
3	10	20
4	10	20
Total	42	85
Maximum marks (Excluding Bonus questions)		60

STOE-3: Applied Statistics

Course Objectives

1. To enable the students to use statistical tools in finance, industries, population studies and health sciences.
2. To acquire knowledge about sampling methods for surveys.

Course Pre-requisite: II PUC or equivalents with Mathematics

Course Outcomes

Upon successful completion of this course, the student will be able to:

1. CO1: Understand the Price and Quantity Index numbers and their different measures, understand the applicability of cost-of-living Index number.
2. CO2: Know the components and Need for Time series, understand the different methods of studying trend and Seasonal Index.
3. CO3: Study the concept of vital statistics, sources of data, different measures of Fertility and Mortality, Understand the Growth rates- GRR and NRR and their interpretations.
4. CO4: Know the concept of Population, Sample, Sampling unit, sampling design, sampling frame, sampling scheme, need for sampling, apply the different sampling methods for designing and selecting a sample from a population, explain sampling and non-sampling errors.
5. CO5: Describe the philosophy of statistical quality control tools as well as their usefulness in industry and hence develop quality control tools in a given situation.

Pedagogy

The course is taught using traditional chalk and talk method using problem solving through examples and exercises. Students are encouraged to use resources available on open sources.

Unit 1: Economic Statistics 12 Hours

Index numbers: Definition, Criteria for a good index number, different types of index numbers. Construction of index numbers of prices and quantities, consumer price index number. Uses and limitations of index numbers. Consumer price index number: construction of consumer price index numbers. Applications of consumer price index numbers

Time Series Analysis: Components of time series, Decomposition of time series- Additive and multiplicative model with their merits and demerits, Illustrations of time series. Measurement of trend by method of free-hand curve, method of semi-averages and method of least squares (linear). Measurement of seasonal variations by method of ratio to trend.

Unit 2: Vital Statistics 10 Hours

Sources of demographic data, errors in data.

Measurement of mortality: crude death rate, specific death rates, and standardized death rates, infant mortality rate, maternal mortality rate, neo natal mortality rates, merits and demerits and comparisons of various mortality rates.

Measurement of Fertility and Reproduction: Fecundity, fertility, measurement of fertility, crude birth rate, general fertility rate, age specific fertility rate and total fertility rates, merits and demerits of each measure of fertility, comparative study of these measures of fertility, Growth rates: Gross reproduction rate and Net reproduction rates.

Unit 3: Sampling Theory 10 Hours

Population and Sample. Need for sampling, Complete Enumeration versus Sample Surveys, Merits and Demerits, non – Probability and Probability Sampling, Need and illustrations. Use of random numbers, Principal steps in sample survey. Requisites of a good questionnaire. Pilot surveys, Sampling and non – sampling errors, Description of SRS, simple random sampling with and without replacement procedures, Merits and demerits of Simple random sampling. Stratified Sampling: Need for stratification, stratifying factors, Merits and demerits of stratified random sampling. Systematic random sampling procedure of obtaining sample, Merits and demerits of systematic random sampling.

Unit 4: Statistical Quality Control 10 Hours

Concept of quality and its management, dimensions of quality, 7 Q C tools, DMAIC principles. Causes of variations in quality: chance and assignable. General theory of control charts, Control charts for variables: X- bar and R-charts. Control charts for attributes: p and c-charts. Acceptance Sampling Plans (Product control): Basic terminologies: AQL, LTPD, AOQ, AOQL, ASN, OC curve, producer's risk, and consumer's risk. Concepts of Single sampling plan, double sampling plan.

References

1. J. Medhi (1992) Statistical Methods. New Age International (P) Ltd. New Delhi.
2. M.N. Das (1993) Statistical Methods and Concepts. Wiley Eastern Ltd.
3. Irwin Miller, John E Freund and Richard A Johnson (1992) Probability and Statistics for Engineers. Prentice Hall of India New Delhi.
4. D.C. Montgomery (1996) Introduction to Statistical Quality Control.
5. Cochran, W G. (1984) Sampling Techniques, Wiley Eastern, New Delhi.
6. Mukhopadhyay P (1998) Theory and Methods of Survey Sampling. Prentice Hall of India.
7. Mukhopadhyay P. (2011): Applied Statistics, 2nd ed. Revised reprint, Books and Allied
8. Kendall M.G. (1976): Time Series, Charles Griffin.
9. Chatfield C. (1980): The Analysis of Time Series –An Introduction, Chapman & Hall.

MODEL BLUEPRINT		
Code number:		**STOE – 3**
Title of the paper:		**Applied Statistics**
Unit no	**No of hrs**	**Total marks**
1	12	25
2	10	20
3	10	20
4	10	20
Total	42	85
Maximum marks (Excluding Bonus Question)		60

STOE – 4: Biostatistics

Course Objectives

1. To enable the students to identify the variables of biological studies and explore the tools of classification and presentation.
2. To study the probability notion, models and their applications in the study of biological phenomenon.
3. To acquire knowledge on sampling distribution and testing of hypotheses.

Course Pre-requisite: II PUC or equivalents with Mathematics

Course Outcomes

After studying the course, the student will be able to apply statistical tools and techniques in data analysis of biological sciences.

Pedagogy

The course is taught using traditional chalk and talk method using problem solving through examples and exercises. Students are encouraged to use resources available on open sources.

Unit 1: Introduction to Bio-Statistics 10 Hours

Definition and scope of Statistics. Scales of Measurement: nominal, ordinal, interval and ratio. Collection, classification and tabulation of data, construction of frequency table for grouped and ungrouped data, graphical representation of data by Histogram, Polygon, Ogive curves and Bar and Pie diagram.

Unit 2: Descriptive Statistics 12 Hours

Measures of Central Tendency: Arithmetic mean, Median and Mode- definition, properties, merits and limitations.
Measures of Dispersion: Range, Standard deviation and Coefficient of Variation.
Correlation and Regression Analysis: Relation between two variables, definition of correlation, types of correlation, Scatter diagram, Karl-Pearson's coefficient of linear correlation and its properties, Spearman's Rank Correlation coefficient.
Regression- Simple linear regression, fitting of regression equations by method of Least Squares, linear regression coefficients and their properties.

Unit 3: Probability and Distributions 10 Hours

Probability: Random experiment, sample space, events-mutually exclusive and exhaustive events. Classical, statistical and axiomatic definitions of probability, addition and multiplication theorems, Bayes' theorem (only statements).
Discrete and continuous random variables, probability mass and density functions, distribution functions, expectation of a random variable.
Standard univariate distributions: Binomial, Poisson and Normal distributions (Elementary properties and applications only).

Unit 4: Sampling Distributions and Statistical Inference 10 Hours

Concepts of random sample and statistic, distribution of sample mean from a normal population, Chi-square, t and F distributions (No derivations) and their applications. Estimation

of population mean, population standard deviation and population proportion from the sample counter parts.

Statistical Hypothesis – null and alternative hypothesis, simple and composite hypothesis. Type I and Type II errors, size, level of significance, power test, critical region, P-value and its interpretation. Test for single mean, equality of two means, single variance, equality of two variances for normal Populations, Test for proportions. Chi-square goodness of fit, One-way ANOVA

References

1. Dutta, N. K. (2004), Fundamentals of Biostatistics, Kanishka Publishers.
2. Gurumani N. (2005), An Introduction to Biostatistics, MJP Publishers.
3. Daniel, W. W. (2007), Biostatistics - A Foundation for Analysis in the Health Sciences, Wiley
4. Rao, K. V. (2007), Biostatistics - A Manual of Statistical Methods for use in Health
5. Nutrition and Anthropology
6. Pagano, M. and Gauvreau, K. (2007), Principles of Biostatistics.
7. Rosner Bernard (2010), Fundamentals of Biostatistics, 6th Edition, Duxbury.

MODEL BLUEPRINT		
Code number:		**STOE – 4**
Title of the paper:		**Biostatistics**
Unit no	**No of hrs**	**Total marks**
1	10	20
2	10	20
3	12	25
4	10	20
Total	42	85
Maximum marks (Excluding Bonus questions)		60

*****