Punyashlok Ahilyadevi Holkar Solapur University, Solapur

NAAC Accredited-2022
'B^{++}' Grade (CGPA 2.96)

Name of the Faculty: Science & Technology

CHOICE BASED CREDIT SYSTEM

Syllabus: STATISTICS

Name of the Course: B.Sc. II (Sem.– III & IV)

(Syllabus to be implemented from June 2023)

PUNYASHLOK AHILYADEVI HOLKAR SOLAPUR UNIVERSITY, SOLAPUR

Proposed Syllabus for B.Sc. Part – II (CBCS Semester Pattern)

STATISTICS

(To be introduced from June 2023)

Introduction

This syllabus of Statistics carries 300 marks. In semester-III, there will be University Theory examination for Paper-V and Paper-VI and in semester-IV, there will be University Theory examination of Paper-VII and Paper-VIII, along with the University Practical examination of Practical Paper-II and Paper-III. The distribution of marks is as below.

Semester No.	Paper No.	Title of the Paper	University Exam.	Internal Exam.	Total	Credits
III	Paper- V	Probability Distributions	40	10	50	4.0
	Paper-VI	Statistical Methods-I	40	10	50	
IV	Paper-VII	Continuous Probability Distributions	40	10	50	4.0
	Paper-VIII	Statistical Methods-II	40	10	50	
Annual Examination	Practical	Paper –II	40	10	50	4.0
		Paper –III	40	10	50	

Note: Nature of internal examination, passing standard, ATKT and the conversion of marks into grades and credits are as per guidelines of Science Faculty Credit and Grading System.\

Teaching Periods:

1) Each theory paper will have three teaching periods per week.
2) Each Practical Paper will have four teaching periods per week per paper per batch of 16 students.

Duration of University Examinations:

1) For theory Paper – V and Paper – VI: Two hours in semester - III.

2) For theory Paper – VII and Paper – VIII: Two hours in semester – IV.

3) For practical paper – II: Four hours for a batch of 16 students annually.

4) For practical paper – III: Four hours for a batch of 16 students annually.

Semester – III

Statistics Paper – V: Probability Distributions

Course Outcomes: Students will acquire knowledge of-

i) Bivariate discrete distributions, independence of bivariate r.v.s, Mathematical expectation of bivariate discrete distribution

ii) Univariate continuous distribution, various measures of continuous r.v.s, and probabilities using its probability distribution.

iii) Bivariate continuous distributions, independence of bivariate r.v.s, Mathematical expectation of bivariate continuous distribution

iv) Transformation of univariate and bivariate continuous random variable.

CONTENTS:

Unit-1: **(15 hrs)**

1.1 Bivariate Discrete Distribution: Definition of bivariate discrete random variable (X,Y) on finite support, Joint p.m.f., and c.d.f., Properties of c.d.f.(without proof), computation of probabilities of events in bivariate probability distribution, marginal and conditional probability distribution. Independence of two discrete r.v.s.

1.2 Mathematical Expectation: Definition of expectation of function of r.v. in bivariate distribution. Theorems on expectation: (i) $E(X+Y) = E(X) + E(Y)$ (ii) $E(XY) = E(X).E(Y)$ when X and Y are independent. Expectation and variance of linear combination of two discrete r.v.s. Definition of conditional mean, conditional variance. Covariance and correlation coefficient. $Cov(aX+bY, cX+dY)$. Distinction between uncorrelated and

independent variables. Proof of the p.g.f. of sum of two independent r.v. as the product of their p.g.f.s.

Unit-2: **(30 hrs.)**

2.1 Continuous Univariate Distributions: Definition of continuous sample space with illustrations, Definition of continuous random variable (r. v.), probability density function (p. d. f.), cumulative distribution function (c. d. f.) and its properties.

Expectation of a r. v., expectation of a function of r. v., mean, median, mode, quartiles, variance, harmonic mean, geometric mean, raw and central moments, problems.

Moments generating function (m. g. f.): definition and properties

(i) Standardization property $M_X(0) = 1$, (ii) Effect of change of origin and scale,

(iii) Uniqueness property of m. g. f., (if exists, statement only).

Generation of raw and central moments.

Cumulant generating function (c. g. f.): definition, relations between cumulants and central moments (up to order four). Problems.

2.2 Continuous Bivariate Distributions: Definition of bivariate continuous r. v. (X, Y), Joint p. d. f., c. d. f with properties, marginal and conditional distribution, independence of r.v.s, evaluation of probabilities of various regions bounded by straight lines.

Expectation of function of r.v.s, mean, variance, covariance, correlation coefficient, conditional expectation, regression as conditional expectation if it is linear function of other variable and conditional variance, proof of:

i) $E(X \pm Y) = E(X) \pm E(Y)$, ii) $E[E(X/Y)] = E(X)$.

Independence : If X and Y are independent r.v.s, then

(i) $E(XY) = E(X)\,E(Y)$,

(ii) $M_{x+y}(t) = M_x(t) \times M_y(t)$.

(iii) $M_{x,y}(t_1, t_2) = M_x(t_1, 0) \times M_y(0, t_2)$.

(iv) $M_{x+y}(t, t) = M_x(t, 0) \times M_y(0, t) = M_{x+y}(t)$.

Numerical Problems.

2.3: Transformations of a continuous random variable: Transformation of a univariate continuous r. v.: Distribution of $Y = g(X)$, where g is monotonic or non-monotonic function using the following methods:

(i) Jacobian of transformation, (ii) Distribution function and (iii) m.g.f.methods.

Transformation of a continuous bivariate r.v.s: Distribution of a bivariate r.v.s using Jacobin of transformation. Problems.

References and Recommended Readings

1. Parimal Mukhopadhyaya: An Introduction to the Theory of Probability. World Scientific Publishing.

2. Hogg R. V. and Craig A.T.: Introduction to Mathematical Statistics (Third edition),

Macmillan Publishing, New York.

3. Gupta S. C. & Kapoor V.K.: Fundamentals of Mathematical Statistics. Sultan Chand & sons, New Delhi.

4. Goon, A.M., Gupta M.K. and Dasgupta B: Fundamentals of Statistics Vol. I andVol. II World Press, Calcutta.

5. Dr. Kore B. G. and Dr. Dixit P. G.: "Probability Distributions-I", Nirali Prakashan, Pune.

6. Mood A.M., Graybill F.A.: Introduction to theory of Statistics. (Chapter II, IV, V,VII) and Bose D.C. Tata, McGraw Hill, New Delhi. (Third Edition)

7. Walpole R.E. & Mayer R.H.: Probability & Statistics. (Chapter 4, 5, 6, 8, 10) MacMillan Publishing Co. Inc, New York.

Semester – III

Statistics Paper – VI: Statistical Methods-I

Course Outcomes: Students will acquire knowledge of:

i) Sampling theory and its need.

ii) Different methods of sampling and estimators associated with them.

iii) Different rates with respect to population change.

iv) multiple and partial correlation as well as regression for the case of trivariate data.

Unit 1: **(27hrs.)**

1.1 Sampling Theory: Definition of population, sample, parameter, statistic, sample survey, census survey. Advantages of sample survey over census survey, estimator,

unbiased estimator. Methods of sampling: i) Deliberate (purposive) sampling ii) probability sampling and iii) Mixed sampling.

Simple random sampling without replacement (SRSWOR): Some results :

i) Probability of a specified unit being selected in sample at any given draw is equal to $\frac{1}{N}$

ii) Probability of a specific unit being included in the sample is $\frac{n}{N}$.

iii) Probability of drawing a specific sample of size 'n' from a population of size N units is $\frac{1}{\binom{N}{n}}$

iv) $E(\bar{y}_n) = \bar{Y}_N$

v) $E(N\bar{y}_n) = \sum Y_i$ = Population total

vi) $Var(\bar{y}_n) = \frac{(N-n)}{Nn} S^2$

vii) $E(s^2) = S^2$

viii) Estimated variance of sample mean

Simple random sampling with replacement (SRSWR): Some results :

i) $E(\bar{y}_n) = \bar{Y}_N$

ii) $E(N\bar{y}_n) = \sum Y_i$ = Population total

iii) $Var(\bar{y}_n) = \frac{(N-1)}{Nn} S^2$

iv) Estimated variance of sample mean

Standard error of sample means, comparison ofSRSWR and SRSWOR.

1.2 Demography: Introduction, vital events and need of vital statistics, Measures of fertility: Crude Birth Rate (CBR), Age Specific Fertility Rate (ASFR), General Fertility Rate (GFR), Total Fertility Rate (TFR), Measures of reproduction: Gross Reproduction rate (GRR), Net Reproduction Rate (NRR), Measures of mortality: Crude death rate (CDR), Specific Death Rate (SDR) by i) Direct method ii) Indirect method, Standardized Death Rate (STDR), Population projection at time t, Life Table - construction and its applications in insurance, Use and Applications

Unit 2: (18 hrs)

2.1 Multiple Linear Regression (for trivariate data only): Concept of multiple linear regression, Plane of regression, Yule's notation, correlation matrix. Fitting of regression plane by method of least squares, definition of partial regression coefficients and their

interpretation. Residual: definition, order, properties, derivation of mean and variance, covariance between residuals.

2.2 Multiple and Partial Correlation (for trivariate data only): Concept of multiple correlation. Definition of multiple correlation coefficient $R_{i.jk}$, derivation of formula for multiple correlation coefficient. Properties of multiple correlation coefficient :

i) $0 \leq R_{i.jk} \leq 1$, (ii) $R_{i.jk} > \max\{|r_{ij}|, |r_{ik}|, |r_{ij.k}|, |r_{ik.j}|\}$

(iii) $R_{i.jk} \geq |r_{ik}|$ i = j =k = 1, 2, 3. i≠j, i≠k. Interpretation of $R_{i.jk}$ =1, $R_{i.jk}$ = 0, coefficient of multiple determination $R^2_{ij.k}$

Concept of partial correlation. Definition of partial correlation coefficient $r_{ij.k}$. Properties of partial correlation coefficient; (i) $-1 \leq r_{ij.k} \leq 1$, (ii) $b_{ij.k} \times b_{ji.k} = r^2_{ij.k}$.

Numerical problems.

References and Recommended Readings

1. Cochran, W.G: Sampling Techniques, Wiley Eastern Ltd., New Delhi.

2. Des Raj: Sampling Theory.

3. Gupta S. C. and Kapoor V. K., "Fundamentals of Applied Statistics", Sultan and Chand, (2010).

4. Dr. Kore B. G. and Dr. Dixit P. G.: "Statistical Methods-I", Nirali Prakashan, Pune.

5. Mukhopadhay, Parimal: Theory and Methods of Survey Sampling, Prentice Hall.

6. Montgomery D. C. (2009). "Introduction to quality Control", Jon Wiley and sons.

7. Sukhatme, P.V. and Sukhatme, B.V.: Sampling Theory of Surveys with Applications, Indian Society of Agricultural Statistics, New Delhi

Semester – IV

Statistics Paper – VII: Continuous Probability Distributions

Course Outcomes: Students will acquire knowledge of:

i) Uniform, exponential and beta distribution along with their basic properties.

ii) Normal and standard normal distribution and its wide applications

iii) Exact sampling distributions that are used in inferential procedures.

Unit-1: **(22hrs.)**

1.1 Uniform Distribution: p. d. f. of Uniform distribution

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

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

Notation X~U(a, b), c.d.f., m.g.f., mean, variance, moments, β_1, β_2, γ_1 and γ_2 coefficients. Distribution of (i)(X-a) / (b-a), ii) (b-X) / (b-a), (iii) Y = F(x) where F(x) is c.d.f. of any continuous r.v. Problems.

1.2 Exponential distribution: (one parameter): p.d.f.

$$f(x) = \theta e^{-\theta x} x \geq 0, \theta > 0$$

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

Notation X ~ Exp(θ), c.d.f., m.g.f.,c.g.f., mean, variance, C.V., moments, β_1, β_2, γ_1 and γ_2 coefficients, median, quartiles, lack of memory property, distribution of Y = -($\frac{1}{\theta}$)log(X) where X~U (0, 1). Exponential distribution with scale and location parameters.

1.3Gamma distribution: p.d.f. (two parameters)

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

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

Notation $X \sim G(\alpha, \lambda)$, special cases i) $\alpha = 1, ii)\ \lambda = 1$, m.g.f., c.g.f., mean mode, variance, moments, β_1, β_2, γ_1 and γ_2 coefficients, additive property, distribution of sum ofi. i. d. exponential variates.

1.4 Beta distribution of first kind: p. d. f.

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

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

Notation $X \sim \beta_1(m,n)$ symmetry around mean when m = n, mean, harmonic mean, mode, variance, Uniform distribution as a particular case when m = n = 1, distribution of(1-X).

1.5 Beta distribution of second kind: p. d. f.

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

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

Notation $X \sim \beta_2(m,n)$ mean, harmonic mean, mode, variance, distribution of $\left(\frac{1}{X}\right)$, relation between beta distribution of firstkind and second kind, distribution of X+Y, $\left(\frac{X}{Y}\right)$

and $\left(\frac{X}{X+Y}\right)$ where X and Y areindependent gamma variate.

Unit-2: **(23 hrs.)**

2.1 Normal Distribution: p. d. f.

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

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

Notation $X \sim N(\mu, \sigma^2)$ definition of standard normal distribution, properties of normal curve, m. g. f., c.g.f., mean, variance, median, mode, mean deviation, quartiles, point of inflexion, moments, recurrence relation for central moments,β_1, β_2, γ_1 and γ_2 coefficients, standard normal distribution, additive property, distribution of X^2 if X~N(0,1), distribution of aX+bY+c when X and Y are independent normal r.v.s. Normal as a limiting case of i) Binomial ii) Poisson (without proof), illustrations of use of normal distribution in various fields.

2.2 Exact Sampling Distributions

Chi-Square distribution: Definition of chi square variate as a sum of square of n i. i. d. standard normal variates,derivation of p.d.f. of chi square distribution with n degrees of freedom using m.g.f., c.g.f., mean, mode, variance, moments,β_1, β_2, γ_1 and γ_2 coefficients,

m. g. f. additiveproperty, relation with gamma distribution. Normal approximation to chi-square distribution using central limit theorem.

Student's t- distribution: Definition of student's t variate in the form $t = \frac{U}{\sqrt{\frac{\chi^2}{n}}}$ where U~N(0,1) and χ^2 is chi-square variate with n d. f. Derivation of p.d.f., mean,mode, variance, moments, β_1, β_2, γ_1 and γ_2 coefficients.

Snedecor's F distribution: Definition of F variate with n_1, n_2 d.f. as a ratio of two independent chi-square variables divided by their respective degrees of freedom. Derivation of p.d.f., mean, variance and mode. Distribution of $\frac{1}{F}$. Interrelation between t, F and χ^2.

References and Recommended Readings

1. Trivedi R. S.: Probability and Statistics with Reliability and Computer Science Application, Prentice – Hall of India Pvt. Ltd., New Delhi.
2. Parimal Mukhopadhyaya: An Introduction to the Theory of Probability. World Scientific Publishing.
3. Hogg R.V. and Craig A.T.: Introduction to Mathematical Statistics (Third edition), Macmillan Publishing, New York.
4. Goon, A.M., Gupta M.K. and Dasgupta B: Fundamentals of Statistics Vol. I and Vol. II World Press, Calcutta.
5. Gupta S. C. & Kapoor V.K.: Fundamentals of Mathematical Statistics. Sultan Chand & sons, New Delhi.
6. Gupta S. C. & Kapoor V.K.: Applied Statistics. Sultan Chand & sons, New Delhi.
7. Dr. Kore B. G., Dr. Dixit P. G. and Mr. P. S. Kapre: "Probability Distributions-II", Nirali Prakashan, Pune.
8. Mood A.M., Graybill F.A.: Introduction to theory of Statistics. (Chapter II, IV, V, VII) and Boes D.C. Tata, McGraw Hill, New Delhi. (Third Edition)
9. Walpole R.E. & Mayer R.H.: Probability & Statistics. (Chapter 4, 5, 6, 8, 10) MacMillan Publishing Co. Inc, New York

Semester – IV

Statistics Paper – VII: Statistical Methods-II

Course Outcomes: Students will acquire knowledge of:

i) Chebychev's inequality and its application

ii) Time series and its components, measuring the trend ad seasonal components.

iii) The inferential procedure for large as well as small samples.

CONTENTS:

Unit-1: (18 hrs.)

1.1 Chebychev's inequality: Chebychev's inequality for discrete and continuous distribution and problems. Central limit theorem: Statement and proof (based on mgf) for iid r.v.s with finite variance. Problems based on Bernoulli, Binomial, Poisson, Geometric and Chi-Square distributions.

1.2 Time Series: Meaning and need of time series analysis, components of time series: (i) Secular trend (ii) Seasonal Variation (iii) Cyclical Variation (iv) Irregular Variation, Additive and Multiplicative model, utility of time series. 13

Measurement of trend: (i) Moving average method (ii) Least square method.

Measurement of seasonal indices by simple average method.

Unit 2: Testing of Hypothesis: (27 hrs.)

2.1Testing of Hypothesis: Notion of Population, Sample, Parameter, Statistic,Sampling distribution of Statistic, hypothesis, Simple and composite hypothesis, Null and alternative hypothesis, type I and type II errors, Critical region, level ofsignificance, one and two tailed test, power of test. General procedure of testing of hypothesis.

2.2: Large Sample Tests:

a) Tests for means: i) testing of population mean;H_0: $\mu = \mu_0$,

ii) testing equality of population means; H_0: $\mu_1 = \mu_2$

b) Tests forProportion: i) testing of population Proportion; H_0: $P = P_0$

ii) testing equality ofpopulation Proportion; H_0:$P_1 = P_2$

c) test for population correlation: i) H_0: $\rho = \rho_0$ii) H_0: $\rho_1 = \rho_2$ (by Z-transformation)

2.3: Small sample tests: Definition of Fisher's t- variate,

t - test: a) test for means: i) H_0: $\mu = \mu_0$, ii) H_0: $\mu_1 = \mu_2$, ($\sigma_1 = \sigma_2$), iii) Paired t- test

χ^2 - test: i) test for population variance $H_0: \sigma^2 = \sigma_0^2$ (Mean known and unknown),

ii) test for goodness of fit, iii) test for independence of attributes; a) m x n contingency table, b) 2 x 2 contingency table, test statistic with proof. Yate's correction for continuity.

F – test: test for equality of two population variances $H_0: \sigma_1^2 = \sigma_2^2$

References and Recommended Readings

1. Chatfield C. (2004), "The Analysis of Time Series –An Introduction", Chapman & Hall.

2. Gupta S. C. & Kapoor V. K., "Fundamentals of Applied Statistics", Sultan Chand & Sons, New Delhi.

3. Kendall M.G. (1978), "Time Series", Charles Griffin.

4. Dr. Kore B. G. and Dr. Dixit P. G.: "Statistical Methods-II", Nirali Prakashan, Pune.

5. Snedecor G.W. and Cochran W. G. "Statistical Methods", Iowa State University Press.

Practical Course at B.Sc. Part- II

Objectives:

By the end of course students are expected to:

i. Compute probabilities of standard probability distributions.

ii. Compute the expected frequency and test the goodness of fit.

iii. Drawing random samples from standard probability distributions.

iv. Compute the multiple and partial correlation coefficients.

v . Selection of samples by SRS.

vi. Computation and interpretation of vital statistics.

vii. Construction of control chart.

viii. Obtain results using soft wares like MS-Excel.

PRACTICAL –II

1.Fitting of Discrete Uniform, Binomial and Hypergeometric distribution and test for goodness of fit.

2. Fitting of Poisson, Geometric and Negative Binomial distribution and test for goodness of fit.

3. Fitting of Continuous Uniform, Exponential and Normal distribution and test for goodness of fit.

4. Fitting of Normal distribution and test for goodness of fit.

5. Model sampling from Discrete Uniform, Binomial and Hypergeometric distribution.

6. Model sampling from Poisson, Geometric and Negative Binomial distribution.

7. Model sampling from Continuous Uniform, Exponential and Normal distribution.

8. Model sampling from Normal distribution.

9. Simple Random Sampling with Replacement

10. Simple Random Sampling without Replacement

11. Demography-I (Mortality Rates)

12. Demography-II (Fertility Rates, Reproduction rates)

13. Multiple regression

14 Multiple correlation

15. Partial correlation.

16. Bivariate Discrete distribution-I. (Marginal and conditional distribution, computation of probabilities of events).

17. Bivariate Discrete distribution-II (Expectation /conditional expectation / variance / conditional variance /covariance / correlation coefficient)

18. Bivariate Continuous distribution-I. (Marginal and conditional distribution, computation of probabilities of events).

19. Bivariate Continuous distribution-II (Expectation /conditional expectation / variance / conditional variance /covariance / correlation coefficient)

20. Transformation of univariate and bivariate continuous distribution.

PRACTICAL –III

1. Fitting of straight lines and second degree curves

2. Fitting of curves of type $Y=ab^x$, $Y=aX^b$ and $Y=ae^{bX}$

3. Application of Chebychev's inequality

4. Application of Central Limit Theorem.

5. Large sample tests for means.

6. Large sample tests for proportions.

7. Tests for population correlation coefficients (Using Fisher's Z transformation)

8. Tests based on Chi-square distribution.(for population variance, for goodness of fit)

9. Tests for independence.

10. Tests based on t distribution ($\mu=\mu_0$, $\mu_1=\mu_2$, paired and unpaired)

11. Tests based on F distribution ($\sigma_1^2 = \sigma_2^2$)

12. Time Series- I (Estimation of trend) i) by moving average ii) by least square method.

13. Time Series- II (Computation of seasonal indices)i) by simple average method

ii) by ratio to moving average.

14. Applications of Exponential distribution.

15. Applications of Normal distribution.

16-20. Project equivalent to 5 practicals.

Notes:

i) Students are allowed to use computer (MS-Excel) for computation in practicals.
ii) Computer printouts are to be attached to the journal.
iii) Student must complete the entire practical to the satisfaction of the teacher concerned.
iv) Student must produce the laboratory journal along with the completion certificate duly signed by Head of Department at the time of practical examination.

Laboratory requirements:

Laboratory should be well equipped with sufficient number of computers along with necessary software, printers and UPS.

Nature of practical question paper

a) Each practical paper is of 40 marks containing four questions each of 15 marks and student has to solve any two questions using any type of calculator or computer software.

b) Five marks are reserved for viva and five marks are reserved for journal for each paper.

c) If computer software is used, it should be demonstrated online to the examiner.

d) The duration of practical examination will be of three hours which includes computation work, viva and online demonstration.

Equivalence for Theory Papers

Old Syllabus		**New Syllabus**	
Semester No. Paper No.	**Title of the Paper**	**Semester No. Paper No.**	**Title of the Paper**
Sem. III Paper V	Probability Distributions-I	Sem. III Paper V	Probability Distributions
Sem. III Paper VI	Statistical Methods	Sem. III Paper VI	Statistical Methods-I
Sem. IV Paper VII	Probability Distributions-II	Sem. IV Paper VII	Continuous Probability Distributions
Sem. IV Paper VIII	Applied Statistics	Sem. IV Paper VIII	Statistical Methods-II