Anekant Education Society's
Tuljaram Chaturchand College of Arts, Science and Commerce, Baramati
Autonomous

Course Structure for F. Y. B. Sc. (Computer Science) STATISTICS

Semester	Paper Code	Title of Paper	No. of Credits
I	UCSST111	Statistical Methods -I	2
	UCSST112	Probability and Some Discrete Probability Distributions	2
	UCSST113	Practical-I	2
II	UCSST121	Statistical Methods- II	2
	UCSST122	Continuous Probability Distributions and Testing of Hypothesis	2
	UCSST123	Practical-II	2

SYLLABUS (CBCS) FOR F. Y. B. Sc.(CS) STATISTICS
(w. e. from June, 2022)

Name of the Programme : B.Sc. Computer Science
Program Code : USCO
Class : F.Y.B.Sc. (Computer Science)
Semester : II
Course Name : Statistical Methods II
Course Code : UCSST121
No. of lectures : 36
Credit : 2 credits

Course Outcomes:

The students will acquire knowledge about the;

1. Relationship between two variable.(correlation)

2. Concept of regression (Simple and multiple)

3. Forecasting (predicting future values of the time series variable).

4. Basic concept of R-Software

TOPICS/CONTENTS:

UNIT 1: Correlation (For ungrouped data) (07L)

1.1 Concept of bivariate data, scatter diagram, its interpretation, concept of correlation, Positive correlation, negative correlation, zero correlation.

1.2 Karl Pearson's coefficient of correlation, properties of correlation coefficient, Interpretation of correlation coefficient, coefficient of determination with interpretation.

1.3 Spearman's rank correlation coefficient (formula with and without ties).

1.4 Numerical problems

UNIT 2: Regression (for ungrouped data) (10L)

2.1 Concept of linear and nonlinear regression.

2.2 Illustrations, appropriate situations for regression and correlation

2.3 Linear regression: Fitting of both lines of regression using least square method.

2.4 Concept of regression coefficients.

2.5 Properties of regression coefficients : $b_{xy} \cdot b_{yx} = r^2$, $b_{xy} * b_{yx} \leq 1$, $b_{xy} = r (\sigma_x /\sigma_y)$ and $b_{yx} = r (\sigma_y /\sigma_x)$.

2.6 Nonlinear regression models: Second degree curve, exponential curves of the type $Y=ab^x$ and $Y=ax^b$.

2.7 Numerical problems related to real life situations

UNIT3: Multiple Regression and Multiple, partial Correlation (For Trivariate Data)   (8L)

3.1 Concept of multiple regressions, Yule's Notations.

3.2 Fitting of multiple regression planes.[Derivation of equation to the plane of regression of X_1 on X_2 and X_3 is expected. Remaining two equations to be written analogously.]

3.3 Concept of partial regression coefficients, interpretations.

3.4 Concept of multiple correlation: Definition of multiple correlation coefficient and its formula.

3.5 Concept of partial correlation. Definition of partial correlation coefficient and Its formula.

UNIT4: Time series   (6L)

4.1 Meaning and utility

4.2 Components of time series

4.3 Additive and multiplicative models

4.4 Methods of estimating trend, moving average method, least squares method and exponential smoothing method(with graph and interpretation).

4.5 Numerical problems related to real life situations

Unit 5 Fundamentals of R-Software   (5L)

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

5.2 Vectors and vector arithmetic

a) Creation of vectors using functions c, seq, rep

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

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

d) Accessing vectors

5.3 Data frames: creation using data. Frame, subset and transform commands, Resident data sets: Accession and summary,p, q, d, r functions.

References:

1 Introduction to Linear Regression Analysis, Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, Wiley

2 Time Series Methods, Brockwell and Davis, Springer, 2006.

3 Time Series Analysis, 4th Edition, Box and Jenkin, Wiley, 2008.

4 Fundamentals of Applied Statistics(3rd Edition), Gupta and Kapoor, S. Chand and Sons,
New Delhi, 1987.

5 Fundamentals of Statistics, Vol. 1, Sixth Revised Edition, Goon, A. M., Gupta, M. K. and Dasgupta, B. (1983). The World Press Pvt. Ltd., Calcutta

6 Statistical Computing Using R- Software, Vishwas R. Pawgi, Third Edition

Name of the Programme : B.Sc. Computer Science
Program Code : USCO
Class : F.Y.B. Sc. (Computer Science)
Semester : II
Course Name : Continuous Probability Distributions
and Testing of Hypothesis
Course Code : UCSST122
No. of lectures : 36
Credit : 2 credits
Course Outcomes:

Students are expected to be able to,

1. Understand standard continuous probability distributions and their application.

2. Apply the idea of estimation and Tests of Hypothesis.

3. Apply testing of hypothesis in real life situations.

4. Use R-Software for computation of Probability and testing of Hypothesis.

TOPICS/CONTENTS:

UNIT 1: Standard Continuous Probability Distributions   (8L)

 1.1 Uniform Distribution: statement of p.d.f., mean, variance, nature of probability curve.

     Theorem (without proof): The distribution function of any continuous r.v. if it is

     Invertible follows U(0, 1) distribution

 1.2 Exponential Distribution: statement of p.d.f. of the form $(x) = (1/\theta) e(-x/\theta)$, mean,

    Variance, nature of probability curve, lack of memory property.(with proof)

 1.3 Normal Distribution: statement of p.d.f., identification of parameters, nature of

     Probability density curve, standard normal distribution, symmetry, distribution of

     aX+b, aX+bY+c where X and Y are independent normal variables, computations of

     Probabilities using normal probability table, normal approximation to binomial and

     Poisson distribution, central limit theorem (statement only), normal probability plot.

 1.4 Numerical problems related to real life situations.

UNIT 2: Statistical Inference: The idea of estimation and Tests of Hypothesis
(4L)
 2.1 Concepts of population and sample.

 2.2 Definitions: random sample from a probability distribution, parameter, statistic,

     Standard error of estimator.

2.3 Concept of null hypothesis and alternative hypothesis (Research hypothesis), critical region, level of significance, type I and type II error, one sided and two sided tests, test of hypothesis, p-value.

UNIT 3: Parametric Tests (18L)

3.1   Large Sample Tests

3.1.1   Ho: $\mu = \mu o$ Vs H_1: $\mu \neq \mu o$, $\mu < \mu o$, $\mu > \mu o$ (One sided and two sided tests)

3.1.2   Ho: $\mu_1 = \mu_2$ Vs H_1: $\mu_1 \neq \mu_2$, $\mu_1 < \mu_2$, $\mu_1 > \mu_2$ (One sided and two sided tests)

3.1.3   Ho: $P = Po$ Vs H_1: $P \neq Po$, $P < Po$, $P > Po$ (One sided and two sided tests)

3.1.4   Ho: $P_1 = P_2$ Vs H_1: $P_1 \neq P_2$, $P_1 < P_2$, $P_1 > P_2$ (One sided and two sided tests)

3.1.5   Numerical problems related to real life situations.

3.2 Test based on F- distribution

3.2.1 F-test for testing significance of equality of two population variances.

3.3 Tests based on t – distribution

3.3.1 Ho: $\mu = \mu o$ Vs H_1: $\mu \neq \mu o$, $\mu < \mu o$, $\mu > \mu o$ (One sided and two sided tests)

3.3.2 Ho: $\mu_1 = \mu_2$ Vs H_1: $\mu_1 \neq \mu_2$, $\mu_1 < \mu_2$, $\mu_1 > \mu_2$ (One sided and two sided tests)

3.3.3 Paired t-test.

3.4 Tests based on Chi square distribution

3.4.1 Chi-square test for goodness of fit

3.4.2 Test for independence of attributes (mxn and 2x2)

3.5 Numerical problems related to real life situations.

UNIT 4: Use of R-Software in statistical inference (6L)

4.1 Probability computation using R.

4 .2 Model sampling from standard Continuous distributions

4.3 Tests of hypothesis using R.

References

1. A First course in Probability, Sheldon Ross. Pearson Education Inc.
2. Statistical Methods (An Introductory Text), Medhi J. 1992, New Age International.
3. Modern Elementary Statistics, Freund J.E. 2005, Pearson Publication.
4.Probability, Statistics, Design of Experiments and Queuing Theory with Applications of Computer Science, Trivedi K.S. 2001, Prentice Hall of India, New Delhi.
5.Gupta S. C. and Kapoor V. K.1987 Fundamentals of Mathematical Statistics(3rd Edition) S.Chand and Sons, New Delhi.
6. Mukhopadhyay P. 2015, Mathematical Statistics (3rd Edition), Books And Allied (P), Ltd.
7. Programmed Statistics, B.L. Agarwal, New Age International Publishers.
8. Common Statistical Tests Kulkarni M.B., Ghatpande, S.B.,Gore S.D.1999 Satyajeet Prakashan,
9. Statistical Computing Using R- Software., Vishwas R. Pawgi, Third Edition

SYLLABUS (CBCS) FOR F. Y. B. Sc.(CS) STATISTICS
(w. e. from June, 2022)

Name of the Programme **:** B.Sc. Computer Science
Program Code **:** USCO
Class **:** F.Y.B. Sc. (Computer Science)
Semester **:** II
Course Name **:** Practical – II
Course Code **:** UCSST123
No. of lectures **:** 36
Credit **:** 2 credits

Course Outcomes:

At the end of this course students are expected to be able to

1. Use statistical tools like Fitting of linear regression model (non-linear regression, Model sampling with the help of R-Software.
2. Use statistical tools like correlation, F test, t test, χ^2 test, and Time Series with the help of MS-Excel.

Sr. No.	Title of the Practical
1	Linear correlation and regression (use of scatter plot for explaining the linear Relationship between two variables) using Excel.
2	Fitting of non-linear regression. (use of scatter plot for explaining the nonlinear relationship between two variables) using R-Software.
3	Fitting of normal distribution and computation of expected frequencies.
4	Fitting of the linear regression model (Simple and Multiple) and non-linear Regression models and finding the best fit by using R-Software..
5	Model sampling from continuous uniform, exponential and normal Distributions using R-Software.
6	Large sample tests.
7	F test, t test, χ^2 test using EXCEL (one problem each with equal and unequal variance)(χ^2 test – for the goodness of fit-use fitted problems of Binomial, Poisson and Normal distribution in previous practical problems)
8	Time Series- Estimation of trend by using the method of moving averages using Excel.
9	Write a report on the application of some statistical techniques in the field of computers. (Individual activity)
10	Project (Part-II) - Analysis of data collected in semester - I