Probability and Statistics

Probability and Statistics

Authors

Dr. J. Suresh Goud
Associate Professor
Department of Freshman Engineering
Institute of Aeronautical Engineering College
Dundigal, Hyderabad - 500043

Mr. Ch. Soma Shekar
Assistant Professor
Deparment of Freshman Engineering
Institute of Aeronautical Engineering College
Dundigal, Hyderabd - 500043

AMARAVATI PUBLICATIONS

The Temple of Education....

All rights reserved. No part of this publication which is material protected by this copyright notice may be reproduced or transmitted or utilized or stored in any form or by any means now known or hereinafter invented, electronic, digital or mechanical, including photocopying, scanning, recording or by any information storage or retrieval system, without prior written permission from the ***Publisher.***

Information contained in this book has been published by **StudentsHelpline Publishing House (P) Ltd. , Hyderabad** and has been obtained by its Authors from sources believed to be reliable and are correct to the best of their knowledge. However, the Publisher and its Authors shall in no event be liable for any errors, omissions or damages arising out of use of this information and specifically disclaim any implied warranties or merchantability or fitness for any particular use.

M/s Amaravati Publications, Hyderabad

A Part of StudentsHelpline Publishing House (P) Ltd. , Hyderabad

(An ISO 9001 : 2015 Certified Company)

Head Office

326/C, Level - 3, Surneni Nilayam
Near B K Guda Park, S R Nagar, Hyderabad - 500 038, INDIA
P.No:+91 40 23710657, 238000657 Fax: +91 40 23810657

Reg. Off

5-68, Pedda Gorpadu, Pakala, Tirupati, Chittoor - 517 112 AP, INDIA
mail:studentshelpline.in@gmail.com

© Amaravati Publications

First Edition-2020

ISBN: 978-81-941197-0-8

` **370/- Student Edition**

` **470/- Library Edition with HB**

Printed at StudentsHelpline Group, S R Nagar, Hyderabad-38
Published by Surneni Mohan Naidu for Amaravati Publications, Hyderabad - 38

Probability and Statistics

Module-I: Probability and Random Variables

Probability, Conditional Probability, Baye's Theorem, Random Variables: Basic Definitions, Discrete and Continous random variables; Probability Distribution: Probbaility Mass Function and Probability Density Function, Mathematical Expectation.

Module-II: Probability Distribution

Binomial Distribution: Mean and Variances of Binomial Distribution, Recurrence Formula for the Binomial Distribution; Poisson Distribution: Poisson Distribution as a Limiting Case of Binomial Distribution, Mean and Variance of Poisson Distribution, Recurrence Formula for the Poisson Distribution; Normal Distribution: Mean, Variance, Mode, Median, Characteristics of Normal Distribution.

Module-III: Correlation and Regression

Correlation: Karle Pearson's Coefficient of Correlation, Computation of Correlation Coefficient, Rank Correlation, Repeated Ranks, Properties of Correlation.

Regression: Lines of Regression, Regression Coefficient, Properties of Regression Coefficient, Angle between Lines of Regression, Multiple Correlation and Regression.

Module - IV: Test of Hypothesis-I

Sampling: Definitions of population, Sampling, Parameter of Statistics, Standard error, Test of Significance: Null Hypothesis, Alternate Hypothesis, Type I and Type II errors, Critical Region, Confidence Interval, Level of Significance, One Sided Test, Two Sided Test. Large Sample Test: Test of Significance for Single Mean, Test of Significance for Difference between two Sample Means, Test of Significance for Single Proportion and Test of Difference between Proportions

Module - V: Test of Hypothesis-II

Small Sample Tests; Student t-distribution, its properties: Test of Significance difference between Sample Mean and Population Mean, Difference between Means of two Small Samples; Snedecor's F-distribution and its Properties, Test of Equality of two Population Variances; Chi-Square Distribution and its Properties: Test of Equality of two Population Variances, Chisquare Test of Goodness of Fit

Probability and Statistics

Chapter-I: Descriptive Statistics

1.0 Aims and Objectives **1**

1.1 Introduction **1**

1.2 Data Science **2**

1.2.1 Classification of the Science of Statistics **2**

1.3 Meaning of Statistics **3**

1.4 Nature of Statistics **3**

1.4.1 Statistics as an Art **3**

1.4.2 Statistics as a Science **3**

1.5 Functions of Statistics **4**

1.6 Scope, Uses and Importance of Statistics **4**

1.7 Collection of data **4**

1.7.1 Types of Data **5**

1.7.2 Sources of Data **5**

1.8 Primary data **6**

1.8.1 Methods of Collecting Primary Data **6**

1.8.1.1 Direct Personal Investigation **6**

1.8.1.2 Indirect Personal Investigation **6**

1.8.1.3 Information through Correspondents **7**

1.8.1.4 Questionnaire Method **7**

1.9 Secondary Data **8**

1.10 Meaning of Population and Sample **9**

1.11 Variable **10**

1.12 Data Visualization **11**

1.12.1 Characteristics of Graphs **12**

1.12.2 General Rules for Graphic Presentation **12**

1.12.3 Types of Graphs **13**

1.13 Arithmetic Mean Meaning **25**

1.14 Calculation of Simple Arithmetic Mean **25**

1.14.1 Individual Observations **25**

1.14.2 Calculation of Arithmetic Mean: Discrete Series **26**

1.14.3 Calculation of Arithmetic Mean – Continuous Series **26**

1.14.4 Solved Example **27**

1.15 Weighted Arithmetic Mean **28**
1.16 Definition of Median **30**
1.17 Calculation of Median **30**
1.17.1 Individual Observations **30**
1.17.2 Discrete Series **30**
1.17.3 Continuous Series Exclusive Form **30**
1.17.4 Cumulative Series **31**
1.17.5 Calculation of Missing Frequencies when Median is Known **32**
1.17.6 Calculation of Quartiles, Deciles and percentiles in individual series **33**
1.17.7 Calculation of Quartiles, Deciles, Percentiles in Discrete Series **33**
1.17.8 Calculation of Quartiles, Deciles, Percentiles in Continuous Series **34**
1.18 Mode **35**
1.19 Calculation of Mode-Continuous Series **35**
1.20 Measures of Variability **37**
1.21 Skewness **39**
1.21.1 Tests of Skewness **39**
1.21.2 Measures of Skewness **40**
1.22 Kurtosis **43**
1.22.1 Measures of Kurtosis **43**
1.23 Summary **46**
1.24 Review Questions **46**
1.25 Multiple Choice Questions **47**

Chapter-II: CorrelationI

2.0 Aims and Objectives **49**
2.1 Introduction **49**
2.1.1 Bivariate Distribution **49**
2.2 Correlation **50**
2.2.1 Definitions **50**
2.2.2 Degree of Correlation **51**
2.2.3 Interpreting the Coefficient of Correlation (r) **52**
2.2.4 Importance or Significance of the Correlation Analysis **52**
2.3 Types of Correlation **53**
2.4 Methods of Studying Correlation **56**
2.4.1 Graphic Methods **56**
2.4.1.1 Scatter Diagram Method **56**
2.4.1.2 Merits and Demerits of Scatter Diagram **58**
2.4.1.3 Correlation Graph Method **59**
2.5 Coefficient of Correlation **60**

2.6 Karl Pearson Coefficient of Correlation **60**
2.6.1 Methods of Computation **61**
2.6.2 Change of Scale and Origin in the Calculation of 'r' **65**
2.6.3 Correlation of Grouped Data **66**
2.6.4 Properties of Coefficient of Correlation **68**
2.6.5 Probable Error of the Coefficient of Correlation (PEr) **68**
2.6.6 Coefficient of Determination **69**
2.6.7 Merits and Limitations of Karl Pearson's Coefficient of Correlation **70**
2.6.8 Examples **70**
2.7 Spearman's Rank Correlation Coefficient **79**
2.7.1 Definition 79
2.7.2 When to use Rank Correlation Coefficient **80**
2.7.3 Features of Spearman's Correlation Coefficient **80**
2.7.4 Merits and Limitations of Spearman's Rank Correlation Coefficient **82**
2.7.5 Examples **83**
2.8 Repeated Ranks **85**
2.9 Summary **88**
2.10 Review Questions **89**
2.11 Exercises **90**
2.12 Mutiple Choice Questions **92**

Chapter-III: Regression

3.0 Aims and Objectives **95**
3.1 Introduction **95**
3.2 Regression Analysis **95**
3.2.1 Definitions **96**
3.2.2 Types of Regression **96**
3.3 Methods of Studying Regression **99**
3.3.1 Graphic Method **100**
3.3.2 Algebraic Method **102**
3.4 Regression Equations **103**
3.4.1 Method of Deviations from the Means of X and Y **106**
3.4.2 Regression Coefficients **107**
3.4.3 Properties of Regression Coefficients **107**
3.4.4 Examples **108**
3.4.5 Deviations Taken from Assumed Means **113**

3.5 Regression Analysis in Grouped Data **115**
3.6 Solved Examples **117**
3.7 Angle between the Regression Lines **126**
3.7.1 Examples **127**
3.8 Coefficient of Determination **128**
3.9 Limitation of Regression Analysis **129**
3.10 Curve Fitting **129**
3.10.1 Introduction **129**
3.10.2 Method of Least Squares **131**
3.10.3 Fitting a Straight Line by the Method of Least Squares **134**
3.10.4 Fitting a Parabola or a Second Degree Curve by the Method of Least Squares **135**
3.10.5 Fitting an exponential curve $y = ae^{bx}$ by the method of Least Squares **136**
3.10.6 Fitting a curve $y = ab^{x}$ by the method of Least Squares **137**
3.10.7 Fitting a curve $y = ax^{b}$ by the Method of Least Squares **137**
3.10.8 Examples **138**
3.11 Multiple Regression Analysis **144**
3.11.1 Examples **146**
3.12 Multiple Correlation **149**
3.12.1 Partial and Multiple Correlation **150**
3.12.2 Partial Correlation Co-efficient **151**
3.12.3 Multiple Correlation **151**
3.12.4 Coefficient of Multiple Correlation **153**
3.12.5 Coefficient of Multiple Determination **153**
3.13 Summary **155**
3.14 Review Questions **155**
3.15 Exercises **156**
3.16 Multiple Choice Questions **159**

Chapter-IV: Introduction to Probability

4.0 Aims and Objectives **163**
4.1 Introduction **163**
4.2 Basic Concepts OF Probability **164**
4.3 Mathematical, Statistical Definitions of Probability **170**
4.4 Solved Problems **172**
4.5 Axiomatic Approach to Probability **177**

4.6 Theorems on Probability **178**
4.7 Examples **183**
4.8 Conditional Probability **188**
4.8.1 Examples **191**
4.9 Multiplication Theorem of Probability **192**
4.9.1 Examples **198**
4.10 Total probability theorem **205**
4.11 Baye's Theorem **205**
4.12 Examples **206**
4.13 Summary **211**
4.14 Review Questions **211**
4.15 Exercises **212**
4.16 Multiple Choice Questions **214**

Chapter-V: Random Variables

5.0 Aims and Objectives **217**
5.1 Introduction **217**
5.2 Random Variables **218**
5.2.1 Concept of a Random Variable **218**
5.2.2 Examples **219**
5.2.3 Theorems of Random Variables **220**
5.2.4 Discrete and Continous Sample Space **220**
5.2.4.1 Conditions for a Function to be a Random Variable **220**
5.3 Types of Random Variables **221**
5.4 Functions of Random Variables **222**
5.4.1 Theorems **223**
5.5 Probability Distribution Function **225**
5.5.1 Properties of Distribution Function **225**
5.6 Discrete Probability Distribution **226**
5.6.1 Probability Mass Function **227**
5.6.2 Discrete Distribution Function **228**
5.6.3 Mean, Variance and Standard Deviation of a Probability Distribution **229**
5.6.4 Examples **230**
5.7 ContiNous Probability Distribution **239**
5.7.1 Probability Density Function **239**
5.7.2 Continous Distribution Function **241**

5.7.3 Measures of Central Tendency for Continous Probability Distribution **241**
5.7.4 Examples **242**
5.8 Mathematical Expectation **252**
5.8.1 Examples **253**
5.8.2 Variance of Random Variables **255**
5.8.3 Raw Moments and Central Moments **258**
5.8.3.1 r^{th} MOMENT (ABOUT ORIGIN) **259**
5.8.3.2 Central Moments **260**
5.8.3.3 Relation between μ'_r and μ_r **260**
5.8.4 Addition and Multiplication Theorems (Expectation) **261**
5.8.5 Expectation of a Linear Combination of Random Variables **264**
5.8.6 Covariance **264**
5.8.7 Jointly Distributed Random Variables **265**
5.8.8 Examples **269**
5.9 Moment Generating Function **275**
5.9.1 Cumulants **278**
5.9.2 Characteristic Function **279**
5.9.2.1 Properties of Characteristic Function **279**
5.9.2.2 Theorems on Characteristic Function **280**
5.9.3 Examples **281**
5.10 Summary **285**
5.11 Review Questions **285**
5.12 Exercises **286**
5.13 Multiple Choice Questions **289**

Chapter-VI: Probability Distribution

6.0 Aims and Objectives **293**
6.1 Introduction **293**
6.1.1 Bernoulli's Trials **294**
6.2 Binominal Distribution **294**
6.2.1 Mean and Variance of the Binomial Distribution **297**
6.2.2 Moment Generating Function (M.G.F.) , Mean and Variance of a Binomial Distri bution about Origin **297**
6.2.2.1 Moment Generating Function **297**
6.2.2.2 Mean and Variance of Binomial Distribution **298**

6.2.3 First Four Moments about Origin of Binomial Distribution **298**
6.2.4 Central Moments or Moments about Mean of Binomial Distribution **298**
6.2.5 Additive Property of Binomial Distribution **299**
6.2.6 Recurrence Formula for the Central Moments of the Binomial Distribution **300**
6.2.7 Examples **301**
6.3 Poisson Distribution **310**
6.3.1 Poisson Distribution as Limiting Form of Binomial Distribution **311**
6.3.2 Mean and Variance of Poisson Distribution **312**
6.3.3 Moment Generating Function, Mean and Variance of the Poisson Distribution **312**
6.3.4 Moments About the Origin **313**
6.3.5 Recurrence Formula for the Central moments of the Poisson Distribution **313**
6.3.6 Recurrence relation for Probabilities **314**
6.3.7 Cumulative Function of the Poisson Variate **314**
6.3.8 Reproductive or Additive Property **314**
6.3.9 Examples of the Occurrence of Poisson Distribution in Real Life **315**
6.3.10 Examples **315**
6.4 Normal Distribution **328**
6.4.1 Properties of the Normal Distribution **329**
6.4.2 Normal Distribution as a Limiting Form of Binomial Distribution **335**
6.4.3 Additive Property or Reproductive Property **337**
6.4.4 The Log-Normal Distribution **337**
6.4.5 Chief Characteristics of the Normal Distribution **338**
6.4.6 Importance of Normal Distribution **339**
6.4.7 Examples **340**
6.5 Chebychev's Inequality **351**
6.6 Summary **356**
6.7 Review Questions **356**
6.8 Exercises **357**
6.9 Multiple Choice Questions **361**

Chapter - VII: Sampling Distribution and Estimation

7.0 Aims and Objectives **363**
7.1 Introduction **363**
7.2 Sampling **363**
7.2.1 Census Vs. Sample Survey **364**

7.3 Principles of Sampling **364**
7.3.1 Characteristics of a Good Sample **365**
7.3.2 Advantages of Sampling **365**
7.3.3 Limitations of Sampling **366**
7.4 Types of Sampling **366**
7.5 Parameters and Statistics **366**
7.5.1 Sample Mean and Sample Variance **368**
7.6 Sampling Distribution **369**
7.6.1 Sampling Distribution of the Means (σ known) **371**
7.6.2 Sampling Distribution of Proportions **372**
7.6.3 Sampling Distribution of Differences and Sums **372**
7.7 Central Limit Theorem **374**
7.8 Standard Error of a Statistic **377**
7.9 Sampling Distribution of the Mean (Unknown) **378**
7.10 Sampling Distribution of Variance s^2 **378**
7.10.1 Examples **379**
7.11 Estimation **392**
7.11.1 Introduction **392**
7.11.2 Estimator and Estimate **393**
7.11.3 Element of Estimators **394**
7.11.4 Point Estimation **395**
7.11.4.1 Properties of a Good Estimator **395**
7.11.4.2 More Efficient Unbiased Estimator **403**
7.11.5 Interval Estimation **405**
7.11.5.1 Confidence Interval Estimates of Population Parameters **406**
7.11.5.2 Determination of Proper Sample Size **408**
7.11.5.3 Confidence Intervals of Parameters of Normal Population by Pivot Method **413**
7.11.6 Bayesian Estimation **414**
7.11.7 Examples **414**
7.12 Summary **422**
7.13 Review Questions **422**
7.14 Exercises **424**
7.15 Multiple Choice Questions **426**

Chapter - VIII: Test of Hypothesis - I

8.0 Aims and Objectives **431**

8.1 Introduction **431**

8.2 Concepts of Statistical Hypothesis **432**

8.2.1 Simple and Composite Hypothesis **433**

8.3 Tests of hypothesis and Tests of Significance **433**

8.3.1 Tests of Significance **434**

8.4 Errors in Testing of Hypotheses **436**

8.4.1 Errors in Sampling **436**

8.5 Critical Region and Level of Significance **437**

8.6 One–Tailed and Two-Tailed Tests **438**

8.7 One–Tailed and Two-Tailed Tests **441**

8.8 Procedure for Testing of Hypothesis **442**

8.9 Examples **442**

8.10 Tests of Significance for Large Samples **443**

8.10.1 Test of Significance of Single Means **445**

8.10.1.1 Examples **446**

8.10.2 Test of Significance of the Difference Between the Means of Two Large Samples **452**

8.10.2.1 Examples **453**

8.10.3 Test of Significance for Single Proportion **460**

8.10.3.1 Examples **461**

8.10.4 Test for Significant Difference Between Two Proportions **468**

8.10.4.1 Examples **469**

8.11 Summary **476**

8.12 Review Questions **476**

8.13 Exercises **477**

8.14 Multiple Choice Questions **478**

Chapter - IX: Test of Hypothesis - II

9.0 Aims and Objectives **481**

9.1 Introduction **481**

9.2 Tests of Significance based on χ^2, t and F **482**

9.2.1 Degree of Freedom **482**

9.3 Student's t-Distribution **483**
9.3.1 Properties of t-Distribution **483**
9.3.2 Uses of t-distribution **484**
9.3.3 Students 't' test **484**
9.3.4 Confidence or Fiducial Limits for μ **484**
9.3.5 Critical Values of t and the t-table **485**
9.3.6 Test of Significance of the Difference between Sample Mean and Population Mean **485**
9.3.6.1 Examples **486**
9.3.6.2 Problems related to Student's T-test **495**
9.3.7 Test of Significance of Difference between Means of Two Small Samples Drawn from the Same Normal Population. **500**
9.3.7.1 Examples **501**
9.3.8 Paired Sample t - Test **513**
9.4 F-distribution **514**
9.4.1 Properties of the F-Distribution **517**
9.4.2 Applications **517**
9.4.3 F-test of Significance of the Difference between Population Variances **517**
9.4.4 Examples **518**
9.5 Chi-Square Distribution **528**
9.5.1 Properties of χ^2 - Distribution **529**
9.5.2 Uses of χ^2 - Distribution **530**
9.5.3 Chi-Square (χ^2) Test **530**
9.5.3.1 χ^2 - Test of Goodness of Fit **530**
9.5.3.2 Conditions for the Validity of χ^2 - Test **531**
9.5.3.3 Examples **531**
9.5.4 χ^2 - Test of Independence of Attributes **545**
9.5.4.1 Contingency Table **547**
9.5.4.2 Examples **549**
9.5.5 Chi-Square Test for Population Variance **556**
9.6 Summary **557**
9.7 Review Questions **557**
9.8 Exercises **558**
9.9 Multiple Choice Questions **562**