

COMPUTATIONAL BIostatISTICS

is meant for undergraduate and postgraduate students of life sciences, veterinary, medical sciences (dental, nursing, psychology, nutrition) and the students preparing for civil services examinations conducted by UPSC and NET/JRF test of UGC.

Each chapter deals with an introduction, clear statement of definitions, objective and principles, supplemented by solved examples and ends with a set of exercises. Statistical calculations are mainly based on data obtained from biological experiments. The language of the book has been kept very simple. Collection of multiple choice questions from different examinations is given at the end of each chapter, which makes it the best choice for an in-depth understanding of biostatistics.

The chapter "Computer Programming in Biostatistics" has been added at the end of the book which contains the computer programming of statistical methods in C language, which is the most unique feature of this book.

Contents

- Introduction
- Sampling and Inferences
- Statistical Data
- Graphical Representation of Data
- Measure of Central Tendency
- Measure of Dispersion and Skewness
- Correlation and Regression
- Probability and Distributions
- Tests of Significance
- Chi-square Test
- Vital Statistics (Demography and Measure of Populations)
- Designing and Methodology of Research
- Computer Programming in Biostatistics

Sudhir Kumar Pundir PhD is currently Associate Professor and Head, Department of Mathematics, SD (PG) College, Muzaffarnagar, UP. Formerly, he was Head, Department of Mathematics, Government (PG) College, Rishikesh, UK. Dr Pundir has been a recipient of JRF and SRF of CSIR, New Delhi, during his research work. He has been teaching undergraduate and postgraduate classes for well over two decades. He also organized and attended a number of national and international conferences. Six students have been awarded PhD degree under his supervision. He has to his credit more than 100 books for undergraduate, postgraduate and engineering students which are widely used by the students of various universities.



CBS Publishers & Distributors Pvt Ltd
4819/XI, Prahlad Street, 24 Ansari Road, Daryaganj, New Delhi 110 002, India
E-mail: delhi@cbspd.com, cbspubs@airtelmail.in; Website: www.cbspd.com
New Delhi | Bengaluru | Chennai | Kochi | Kolkata | Lucknow | Mumbai
Hyderabad | Jharkhand | Nagpur | Patna | Pune | Uttarakhand



QR Code to
access online
CBS Catalogue

ISBN: 978-93-5466-172-3



COMPUTATIONAL BIostatISTICS

Sudhir Kumar Pundir



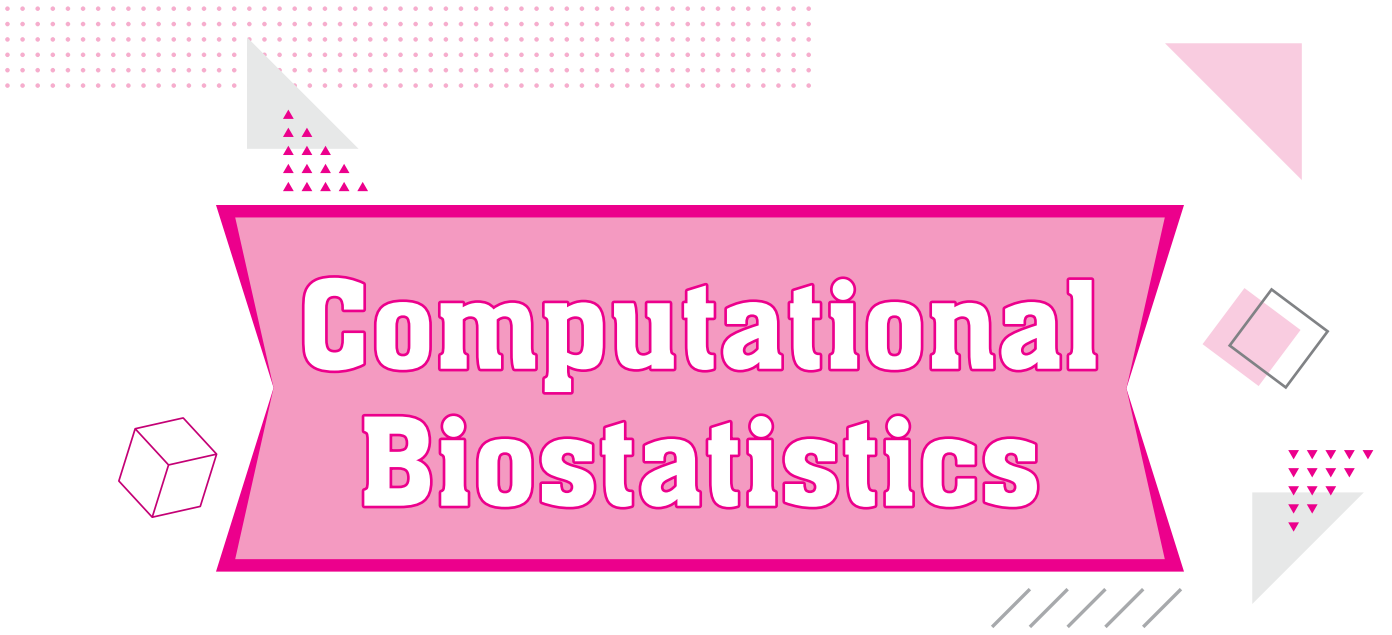
COMPUTATIONAL BIostatISTICS

Useful for _____
Undergraduate and Postgraduate Students of
Medical Sciences, Biomedical Sciences, Biosciences and Life Sciences

Sudhir Kumar Pundir



CBS Publishers & Distributors Pvt Ltd



Computational Biostatistics

Useful For

Undergraduate and Postgraduate Students of
Medical Sciences, Biomedical Sciences, Biosciences and Life Sciences

DR. SUDHIR KUMAR PUNDIR

Msc, MPhil, NET (CSIR), JRF, SRF, PhD

Head

Department of Mathematics

SD (PG) College,
Muzaffarnagar, UP



CBS Publishers & Distributors Pvt Ltd

New Delhi • Bengaluru • Chennai • Kochi • Kolkata • Lucknow • Mumbai
Hyderabad • Jharkhand • Nagpur • Patna • Pune • Uttarakhand



Disclaimer

Science and technology are constantly changing fields. New research and experience broaden the scope of information and knowledge. The author has tried his best in giving information available to him while preparing the material for this book. Although, all efforts have been made to ensure optimum accuracy of the material, yet It is quite possible some errors might have been left uncorrected. The publisher, printer and author will not be held responsible for any inadvertent errors or inaccuracies.

COMPUTATIONAL BIostatISTICS

ISBN : 978-93-5466-172-3

Copyright © Author

First Edition: 2022

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system without permission, in writing, from the author and the publisher.

Published by Satish Kumar Jain and produced by Varun Jain for

CBS Publishers & Distributors Pvt Ltd

4819/XI Prahlad Street, 24 Ansari Road, Daryaganj, New Delhi 110 002, India.

Ph : 23289259, 23266861, 23266867

Website: www.cbspd.com

Fax : 011-23243014

e-mail : delhi@cbspd.com; cbspubs@airtelmail.in

Corporate Office: 204 FIE, Industrial Area, Patparganj, Delhi 110 092

Ph: 011-4934 4934 Fax : 011-4934 4935 e-mail: publishing@cbspd.com, publicity@cbs.com

Branches

- **Bengaluru:** Seema House 2975, 17th Cross, K.R. Road, Banasankari 2nd Stage, Bengaluru 560 070, Karnataka, India
Ph: +91-80-26771678/79 Fax: +91-80-26771680 e-mail: bangalore@cbspd.com
- **Chennai:** 7, Subbaraya Street, Shenoy Nagar, Chennai 600 030, Tamil Nadu, India.
Ph: +91-44-26680620, 26681266 Fax: +91-44-42032115 e-mail: chennai@cbspd.com
- **Kochi:** 42/1325, 36, Power House Road, Opposite KSEB, Kochi-682 018, Kerala, India.
Ph: +91-484-4059061-67 Fax: +91-484-4059065 e-mail: kochi@cbspd.com
- **Kolkata:** 147, Hind Ceramics Compound, 1st Floor, Nilgunj Road, Belghoria, Kolkata-700 056, West Bengal, India
Ph: 9096713055, 7798394118, 9836841399 e-mail: kolkata@cbspd.com
- **Lucknow:** Basement, Khushnuma Complex, 7-Meerabai Marg (Behind Jawahar Bhawan), Lucknow-226 001, Uttar Pradesh, India.
Ph: +0552-4000032 e-mail: tiwari.lucknowi@cbspd.com
- **Mumbai:** PWD Shed, Gala No. 25/26, Ramchandra Bhatt Marg, Next to JJ Hospital Gate No. 2, Opp Union Bank of India, Noorbaug, Mumbai-400009, Maharashtra, India
Ph: +91-22-66661880, 66661889 e-mail: mumbai@cbspd.com

Representatives

- Hyderabad 0-9885175004
- Jharkhand 0-9811541605
- Nagpur 0-9421945513
- Patna 0-9334159340
- Pune 0-9623451994
- Uttarakhand 0-9716462459

Printed at Rashtriya Printers, Dilshad Garden, Delhi, India

Preface

“**COMPUTATIONAL BIostatISTICS**” has assumed tremendous importance with the progress in various disciplines of life sciences. All the universities in India have introduced biostatistics as a subject of study at undergraduate and postgraduate levels. Also, Biostatistics has become indispensable for the research in various fields of life sciences and medical sciences.

Each chapter begins with a brief introduction of the topic and then followed by solved problems, review questions and MCQ. The important facts given to the point at various places of the chapter which is the unique feature of this book. Those researchers as well as students doing dissertation will be benefitted regarding use of different statistical significance tests, which have been dealt lucidly in this book.

I express my gratitude to the authors and publishers of various books I consulted during the preparation of the book.

I wish to sincerely thank Mr SK Jain, Chairman & Managing Director and Mr Varun Jain, Director, CBS Publishers & Distributors, New Delhi, for encouragement and help in bringing out this publication in a present nice form.

My special thanks to Mr YN Arjuna, Senior Vice President, Publishing, Editorial and Publicity and Mrs Ritu Chawla, General Manager, CBS Publishers & Distributors, New Delhi, whose encouragement and unstinted support enabled me to complete the book. I also take this opportunity to express my sincere gratitude to Mr Sunil Dutt, CBS Publishers & Distributors, New Delhi, who gave me the inspiration throughout the preparation of the book. Mr. Suresh Sharma, Mr Ramakant Jha, Mr Anil Rawat and Mr Anurag Singh, CBS, New Delhi, deserve special mention for their kind support and help in this endeavour and unstinted support enabled me to complete my book. Mr Jitesh Ahalawat, M/s Tech Yuva Infostat also deserve special mention for nice type setting..

I must also record my appreciation due to my wife Dr. Rimple, daughter Rijuta and son Shrish for their understanding and love during the long period that I have taken to complete this book.

Above all I am thankful to the Almighty God, without whose grace nothing is possible for any one.

Readers are welcomed to point out errors, if any and send their valuable suggestions for improving the quality of the book.

Dr. Sudhir Kumar Pundir

email : skpundir05@yahoo.co.in

Contents

Preface	V
Ch. 1. INTRODUCTION	1-4
1.1 Introduction	1
1.2 Historical Perspective	1
1.3 Applications of Statistics	1
1.4 Characteristics of Statistics	2
1.5 Limitation of Statistics	2
1.6 Introduction to Bio-statistics	2
1.7 Bio-Statistics	2
1.8 Applications and Uses of Bio-Statistics	3
1.9 Scope of Bio-Statistics	3
Ch. 2. SAMPLING AND INFERENCES	5-10
2.1 Introduction	5
2.2 Sampling	5
2.3 Types of Sampling	5
2.4 Characteristics of a Good Sample Design	7
2.5 Techniques for Random Sampling	7
2.6 Stratified Random Sampling	8
2.7 Systematic Random Sampling	9
2.8 Limitations of Sampling	9
Ch. 3. STATISTICAL DATA	11-32
3.1 Introduction	11
3.2 Sources of Data	12
3.3 Methods of Collecting Primary Data	12
3.4 Schedule and Questionnaire	13
3.5 Methods of Collecting Secondary Data	14
3.6 Classification of Data	14
3.7 Statistical Series	16
3.8 Tabulation of Data	18
3.9 Frequency Distribution	20
3.10 Grouped Frequency Distribution	21
3.11 Cumulative Frequency	24
Ch. 4. GRAPHICAL REPRESENTATION OF DATA	33-54
4.1 Introduction	33
4.2 Types of Diagrams	33
4.3 One-Dimensional Diagram	33
4.4 Two-Dimensional Diagrams	36
4.5 Three Dimensional Diagrams	38
4.6 Pie Diagrams	39
4.7 Cartograms	40
4.8 Pictogram	40
4.9 Histograms	40
4.10 Graphical Representatoin of Data	41
4.11 Graphs of Frequency Distributions	42
4.12 Types of Frequency Curve	45

Ch. 5. MEASURE OF CENTRAL TENDENCY	55-94
5.1 Introduction	55
5.2 Kinds of Statistical Averages	55
5.3 Arithmetic Mean	55
5.4 Combined Mean	66
5.5 Geometric Mean	67
5.6 Harmonic Mean	71
5.7 Median	73
5.8 Quartiles	78
5.9 Quantiles	79
5.10 Decile and Percentile	79
5.11 Mode	84
5.12 Empirical Formula	86
Ch. 6. MEASURE OF DISPERSION AND SKEWNESS	95-136
6.1 Introduction	95
6.2 Measure of Dispersion	95
6.3 Range	96
6.4 Interquartile Range and Semi-Quartile Deviation	98
6.5 Average Deviation or Mean Deviation	101
6.6 Absolute and Relative Measure of Dispersion	106
6.7 Standard Deviation and Root Mean Square Deviation	107
6.8 Short Cut Method to Calculate the Standard Deviation of Discrete Series	111
6.9 Summary of the Mathematical Properties of Standard Deviation	116
6.10 Relation between Different Measures of Central Tendency and Measure of Skewness	118
6.11 Relation between Different Measures of Dispersion	118
6.12 Standard Deviation of Two Combined Sets	123
6.13 Skewness	128
6.14 Measure of Skewness	128
Ch. 7. CORRELATION AND REGRESSION	137-162
7.1 Introduction	137
7.2 Multivariate and Bivariate Data	137
7.3 Correlation	137
7.4 Types of Correlation	138
7.5 Perfect Correlation	138
7.6 Methods for Finding the Correlation	138
7.7 Coefficient of Correlation for Grouped Distribution	144
7.8 Rank Correlation	146
7.9 Regression	150
7.10 Properties of Regression Coefficients	151
7.11 Angle between Two Lines of Regression	151
7.12 Multiple Linear Regression	156
Ch. 8. PROBABILITY AND DISTRIBUTIONS	163-206
8.1 Introduction	163
8.2 Classical (or Prior) Probability	164
8.3 Rules on Probability	168
8.4 Conditional Probability	169
8.5 Dependent and Independent Events	169
8.6 Probability of Happening of at least One of Independent Events	171

8.7	Law of Total Probability and Bayes' Theorem	172
8.8	Random Variables and Probability Distribution	177
8.9	Probability Distributions	177
8.10	Binomial Distribution	180
8.11	Applicability of Binomial Distribution	181
8.12	Moments of the Binomial Distribution	181
8.13	Poisson Distribution	187
8.14	Limiting Form of Binomial Distribution	188
8.15	Moments of Poisson Distribution	188
8.16	Moment Generating Function (M.G.F.) of Poisson Distribution	190
8.17	Cumulant Generating Function of Poisson Distribution	190
8.18	Normal Distribution	194
8.19	Standard Form of the Normal Curve	195
8.20	Properties of the Normal Distribution	195
8.21	Fitting of Normal Distribution	196
8.22	Moment Generating Function of Normal Distribution	197
8.23	Cumulant Generating Function of Normal Distribution	197
Ch. 9.	TESTS OF SIGNIFICANCE	207-252
9.1	Introduction	207
9.2	Types of Population	207
9.3	Statistical Hypothesis	207
9.4	Null and Alternative Hypothesis	208
9.5	Tests of Significance	208
9.6	Level of Significance	208
9.7	Critical Region and Acceptance Region	208
9.8	Type-I And Type-II Errors	209
9.9	Best Critical Region	210
9.10	One Tail and Two Tail Tests	210
9.11	Test of Significance of Large Samples	212
9.12	Standard Error	212
9.13	Probable Error	213
9.14	Test of Significance in Case of Attributes (Large Samples)	213
9.15	Confidence Limits of Unknown P	214
9.16	Test of Significance of Mean (Large Sample of Variables)	215
9.17	Testing the Significance of the Difference between the Means of Two Large Samples	216
9.18	Test of Significance of Difference between the Sample Proportion	219
9.19	Test of Significance of Difference between the Standard Deviation (Large Samples)	220
9.20	Students 't' Test	221
9.21	Paired t-Test for difference of Means	230
9.22	F-Test of Equality of Population Variance	232
9.23	Analysis of Variance (ANOVA)	236
Ch. 10.	CHI-SQUARE TEST	253-270
10.1	Introduction	253
10.2	Chi-Square (χ^2) Test	253
10.3	Applications of χ^2 Distribution	253
10.4	Chi-Square (χ^2) Test for Goodness of Fit	254
10.5	χ^2 -Test for Independence of Attributes	260

Ch. 11. VITAL STATISTICS (DEMOGRAPHY AND MEASURE OF POPULATIONS)	271–284
11.1 Introduction	271
11.2 Demography	271
11.3 Population Census	271
11.4 Vital Biostatistics	271
11.5 Methods of Obtaining Vital Statistics	272
11.6 Measures of Population	272
11.7 Measure of Vital Statistics	274
11.8 Measurement of Mortality	275
11.9 Infant Mortality Rates	276
11.10 Fertility Rates	277
11.11 Life Table	278
11.12 Construction of a Life Table	278
Ch. 12. DESIGNING AND METHODOLOGY OF RESEARCH	285–292
12.1 Introduction	285
12.2 Methodology and Designing	285
12.3 Reserarch Protocols	287
12.4 Some Common Study Designs	288
12.5 Clinical Trial	288
12.6 Research Publications : General Terms	289
Ch. 13. COMPUTER PROGRAMMING IN BIostatISTICS	293–352
13.1 Introduction	293
13.2 Histogram	293
13.3 Arithmetic Mean	294
13.4 Median	298
13.5 Deciles	311
13.6 Percentiles	315
13.7 Mode	318
13.8 Geometric Mean	322
13.9 Harmonic Mean	324
13.10 Mean Deviation	326
13.11 Standard Deviation	328
13.12 Variance and Coefficient of Variation	331
13.13 Correlation Coefficient	335
13.14 Rank Correlation Coefficient	337
13.15 Regression Lines	339
13.16 Partial and Multiple Correlation Coefficient	345
13.17 Binomial Distribution	347
13.18 Poisson Distribution	350
● APPENDIX: SELECTED TABLES	353–368
● BIBLIOGRAPHY	369–370
● INDEX	371–372