# FUNDAMENTALS OF MATHEMATICAL STATISTICS

## (A Modern Approach)

*A Textbook written completely on modern lines for Degree, Honours, Post-graduate Students of all Indian Universities and Indian Civil Services, Indian Statistical Service Examinations.*

**(Contains, besides complete theory, more than 650 fully solved examples and more than 1,500 thought-provoking Problems with Answers, and Objective Type Questions)**

**S.C. GUPTA**
*Reader in Statistics*
*Hindu College,*
*University of Delhi*
*Delhi*

**V.K. KAPOOR**
*Reader in Mathematics*
*Shri Ram College of Commerce*
*University of Delhi*
*Delhi*

**Tenth Revised Edition**
**(Greatly Improved)**

**SULTAN CHAND & SONS**
*Educational Publishers*
**New Delhi**

* First Edition : Sept. 1970
Tenth Revised Edition : August 2000
Reprint : 2002

* Price : **Rs. 210.00**

ISBN 81-7014-791-3



* *Published by* :
**Sultan Chand & Sons**
23, Darya Ganj, New Delhi-110002
Phones : 3277843, 3266105, 3281876

* *Laser typeset by* : **T.P.**

**Printed at:- New A.S. Offset Press Laxmi Nagar Delhi-92**

DEDICATED TO
OUR TEACHER
**PROFESSOR H.C. GUPTA**
WHO INITIATED
THE TEACHING OF
MATHEMATICAL STATISTICS
AT THE
UNIVERSITY OF DELHI

# PREFACE

## TO THE TENTH EDITION

*The book has been revised keeping in mind the comments and suggestions received from the readers. An attempt is made to eliminate the misprints/errors in the last edition. Further suggestions and criticism for the improvement of the book will be most welcome and thankfully acknowledged.*

*August 2000* **S.C. GUPTA**
**V.K. KAPOOR**

## TO THE NINTH EDITION

*The book originally written twenty-four years ago has, during the intervening period, been revised and reprinted several times. The authors have, however, been thinking, for the last few years that the book needed not only a thorough revision but rather a complete rewriting. They now take great pleasure in presenting to the readers the ninth completely revised and enlarged edition of the book. The subject-matter in the whole book has been rewritten in the light of numerous criticisms and suggestions received from the users of the previous editions in India and abroad.*

***Some salient features of the new edition are:***

- *The entire text, especially Chapter 5 (Random Variables), Chapter 6 (Mathematical Expectation), Chapters 7 and 8 (Theoretical Discrete and Continuous Distributions), Chapter 10 (Correlation and Regression), Chapter 15 (Theory of Estimation), has been restructured, rewritten and updated to cater to the revised syllabi of Indian universities, Indian Civil Services and various other competitive examinations.*
- *During the course of rewriting, it has been specially borne in mind to retain all the basic features of the previous editions especially the simplicity of presentation, lucidity of style and analytical approach which have been appreciated by teachers and students all over India and abroad.*
- *A number of typical problems have been added as solved examples in each chapter. These will enable the reader to have a better and thoughtful understanding of the basic concepts of the theory and its various applications.*
- *Several new topics have been added at appropriate places to make the treatment more comprehensive and complete. Some of the obvious ADDITIONS are:*

*§ 8·1.5 Triangular Distribution p. 8·10 to 8·12*
*§ 8·8.3 Logistic Distribution p. 8·92 to 8·95*
*§ 8·10 Remarks 2, Convergence in Distribution of Law p. 8·106*
*§ 8·10.3. Remark 3, Relation between Central Limit Theorem and Weak Law of Large Numbers p. 8·110*
*§ 8·10.4 Cramer's Theorem p . 8·111-8.112, 8·114-8·115 — Example 8.46*

§ 8·14 to § 8·14·6 Order Statistics — Theory, Illustrations and Exercise Set p. 8·136 to 8·151

§ 8·15 Truncated Distributions—with Illustrations p. 8·151 to 8·156

§ 10·6·1 Derivation of Rank Correlation Formula for Tied Ranks p. 10·40-10·41

§ 10·7·1 Lines of Regression—Derivation (Aliter) p. 10·50-10·51. Example 10·21 p. 10·55

§ 10·10·2 Remark to § 10·10·2 — Marginal Distributions of Bivariate Normal Distribution p. 10·88-10·90
Theorem 10·5, p. 10·86. and Theorem 10·6, p. 10·87 on Bivariate Normal Distribution.
Solved Examples 10·31, 10·32, pages 10·96-10·97 on BVN Distribution.
Theorem 13·5 Alternative Proof of Distribution of $(X, s^2)$ using m.g.f. p.13·19 to 13·21

§ 13·11 $\chi^2$-Test for pooling of Probabilities ($P_\lambda$ Test) p. 13·69

§ 15·4·1 Invariance property of Consistent Estimators—Theorem 15·1, pp 15·3

§ 15·4·2 Sufficient Conditions for Consistency—Theorem 15·2, p. 15·3

§ 15·5·5 MVUE : Theorem 15·4, p. 15·12-15·13

§ 15·7 Remark 1. Minimum Variance Bound (MVB), Estimator, p.15·24

§ 15·7·1 Conditions for the equality sign in Cramer-Rao (CR) inequality, p. 15·25 to 15·27

§ 15·8 Complete family of Distributions (with illustrations), p. 15·31 to 15·34
Theorem 15·10 (Blackwellisation), p. 15·36.
Theorems 15·16 and 15·17 on MLE, p. 15·55.

§ 16·5·1 Unbiased Test and Unbiased Critical Region.
Theorem 16·2-pages 16·9-16·10

§ 16·5·2 Optimum Regions and Sufficient Statistics, p.16·10-16·11
Remark to Example 16·6, p. 16·17-16·18 and Remarks 1, 2 to Example 16·7, p. 16·20 to 16·22; Graphical Representation of Critical Regions.

• Exercise sets at the end of each chapter are substantially reorganised. Many new problems are included in the exercise sets. Repetition of questions of the same type (more than what is necessary) has been avoided. Further in the set of exercises, the problems have been carefully arranged and properly graded. More difficult problems are put in the miscellaneous exercise at the end of each chapter.

• Solved examples and unsolved problems in the exercise sets have been drawn from the latest examination papers of various Indian Universities, Indian Civil Services, etc.

• *An attempt has been made to rectify the errors in the previous editions.*

• ***The present edition incorporates modern viewpoints. In fact with the addition of new topics, rewriting and revision of many others and restructuring of exercise sets, altogether a new book, covering the revised syllabi of almost all the Indian universities, is being presented to the reader. It is earnestly hoped that, in the new form, the book will prove of much greater utility to the students as well as teachers of the subject.***

*We express our deep sense of gratitude to our Publishers M/s Sultan Chand & Sons and printers DRO Phototypesetter for their untiring efforts, unfailing courtesy, and co-operation in bringing out the book, in such an elegant form. We are also thankful to our several colleagues, friends and students for their suggestions and encouragement during the preparing of this revised edition.*

*Suggestions and criticism for further improvement of the book as well as intimation of errors and misprints will be most gratefully received and duly acknowledged.*

**S C. GUPTA & V.K. KAPOOR**

## *TO THE FIRST EDITION*

*Although there are a large number of books available covering various aspects in the field of Mathematical Statistics, there is no comprehensive book dealing with the various topics on Mathematical Statistics for the students. The present book is a modest though determined bid to meet the requirements of the students of Mathematical Statistics at Degree, Honours and Post-graduate levels. The book will also be found of use by the students preparing for various competitive examinations. While writing this book our goal has been to present a clear, interesting, systematic and thoroughly teachable treatment of Mathematical Statistics and to provide a textbook which should not only serve as an introduction to the study of Mathematical Statistics but also carry the student on to such a level that he can read with profit the numercus special monographs which are available on the subject. In any branch of Mathematics, it is certainly the teacher who holds the key to successful learning. Our aim in writing this book has been simply to assist the teacher in conveying to the students more effectively a thorough understanding of Mathematical Statistics.*

*The book contains sixteen chapters (equally divided between two volumes). The first chapter is devoted to a concise and logical development of the subject. The second and third chapters deal with the frequency distributions, and measures of average and dispersion. Mathematical treatment has been given to the proofs of various articles included in these chapters in a very logical and simple manner. The theory of probability which has been developed by the application of the set theory has been discussed quite in detail. A large number of theorems have been deduced using the simple tools of set theory. The* *simple applications of probability are also given. The chapters on mathematical expectation and theoretical distributions (discrete as well as continuous) have been written keeping the latest ideas in mind. A new treatment has been given to the chapters on correlation, regression and bivariate normal distribution using the concepts of mathematical expectation. The thirteenth and fourteenth chapters deal mainly with the various sampling distributions and the various tests of significance which can be derived from them. In chapter 15, we have discussed concisely statistical inference (estimation and testing of hypothesis). Abundant material is given in the chapter on finite differences and numerical integration. The whole of the relevant theory is arranged in the form of serialised articles which are concise and to the point without being insufficient. The more difficult sections will, in general, be found towards the end of each chapter. We have tried our best to present the subject so as to be within the easy grasp of students with varying degrees of intellectual attainment.*

*Due care has been taken of the examination needs of the students and, wherever possible, indication of the year, when the articles and problems were set in the examination as been given. While writing this text, we have gone through the syllabi and examination papers of almost all Indian universities where the subject is taught so as to make it as comprehensive as possible. Each chapter contains a large number of carefully graded worked problems mostly drawn from university papers with a view to acquainting the student with the typical questions pertaining to each topic. Furthermore, to assist the student to gain proficiency in the subject, a large number of properly graded problems mainly drawn from examination papers of various universities are given at the end of each chapter. The questions and problems given at the end of each chapter usually require for their solution a thoughtful use of concepts. During the preparation of the text we have gone through a vast body of literature available on the subject, a list of which is given at the end of the book. It is expected that the bibliography given at the end of the book will considerably help those who want to make a detailed study of the subject*

*The lucidity of style and simplicity of expression have been our twin objects to remove the awe which is usually associated with most mathematical and statistical textbooks.*

*While every effort has been made to avoid printing and other mistakes, we crave for the indulgence of the readers for the errors that might have inadvertently crept in. We shall consider our efforts amply rewarded if those for whom the book is intended are benefited by it. Suggestions for the improvement of the book will be highly appreciated and will be duly incorporated.*

**SEPTEMBER 10, 1970** **S.C. GUPTA & V.K. KAPOOR**

# Contents

*Chapter — 3*

**Measures of Dispersion, Skewness and Kurtosis 3·1 — 3·40**



*Chapter — 4*

**Theory of Probability 4·1 — 4·116**

*Chapter 5*

**Random Variables and Distribution Functions 5·1 — 5·82**

*Chapter — 6*

**Mathematical Expectation and Generating Functions** **6·1 — 6·138**

*Chapter - 7*

**Theoretical Discrete Distributions** **7·1 — 7·114**

*Chapter — 11*

**Theory of Attributes** **11·1 — 11·22**



*Chapter — 12*

**Sampling and Large Sample Tests** **12·1 — 12·50**

*Chapter — 15*

**Statistical Inference — I**
**(Theory of Estimation)** **15·1 — 15·92**



*Chapter — 16*

**Statistical Inference — II**
**Testing of Hypothesis, Non-parametric Methods and Sequential Analysis** **16·1 — 16·80**

# Fundamentals of Mathematical Statistics

S.C. GUPTA — *Hindu College, University of Delhi, Delhi*
V.K. KAPOOR — *Shri Ram College of Commerce, University of Delhi, Delhi*

Tenth Edition 2002 | Pages *xx* + 1284 | 22 × 14 cm | ISBN 81-7014-791-3 | **Rs 210.00**

**Special Features**

- Comprehensive and analytical treatment is given of all the topics.
- Difficult mathematical deductions have been treated logically and in a very simple manner.
- It conforms to the latest syllabi of the Degree and Post graduate examinations in Mathematics, Statistics and Economics.

**Contents**

Introduction • Frequency Distribution and Measures of Central Tendency • Measures of Dispersion, Skewness and Kurtosis • Theory of Probability • Random Variables—Distribution Function • Mathematical Expectation, Generating Functions and Law of Large Numbers • Theoretical Discrete Distributions • Theoretical Continuous Distributions • Curve Fitting and Principle of Least Squares • Correlation, Regression, Bivariate Normal Distribution and Partial & Multiple Correlation • Theory of Attributes • Sampling and Large Sample Tests of Mean and Proportion • Sampling Distribution Exact (Chi-square Distribution) • Exact Sampling Distributions (t, F and Z Distributions) • Theory of Estimation • Testing of Hypothesis, Sequential Analysis and Non-parametric Methods.

# Elements of Mathematical Statistics

S.C. GUPTA V.K. KAPOOR

Third Edition 2001 | Pages *xiv* + 489 | ISBN 81-7014-291-1 | **Rs 70.00**

- Prepared specially for B.Sc. students, studying Statistics as subsidiary or ancillary subject.

**Contents**

Introduction—Meaning and Scope • Frequency Distributions and Measures of Central Tendency • Measures of Dispersion, Skewness and Kurtosis • Theory of Probability • Random Variables—Distribution Functions • Mathematical Expectation, Generation Functions and Law of Large Numbers • Theoretical Discrete Distributions • Theoretical Continuous Distributions • Curve Fitting and Principle of Least Squares • Correlation and Regression • Theory of Attributes • Sampling and Large Sample Tests • Chisquare Distribution • Exact, Sampling Distribution • Theory of Estimation • Testing of Hypothesis • Analysis of Variance • Design of Experiments • Design of Sample Surveys • Tables.

# Fundamentals of Applied Statistics

S.C. GUPTA V.K. KAPOOR

Third Edition 2001 | Pages *xvi* + 628 | ISBN 81-7014-151-6 | **Rs 110.00**

**Special Features**

- The book provides comprehensive and exhaustive theoretical discussion.
- All basic concepts have been explained in an easy and understandable manner.
- 125 stimulating problems selected from various university examinations have been solved.
- It conforms to the latest syllabi of B.Sc. (Hons.) and post-graduate examination in Statistics, Agriculture and Economics.

**Contents**

Statistical Quality Control • Analysis of Time Series (Mathematical Treatment) • Index Number • Demand Analysis, Price and Income Elasticity • Analysis of Variance
Design of Experiments, Completely Randomised Design • Randomised Block Design, Latin Square Design • Factorial Designs and Confounding.
Design of Sample Surveys (Mathematical Treatment) • Sample Random Sampling, Stratified Sampling, Systematic Sampling, Multi-stage Sampling • Educational and Psychological Statistics • Vital Statictical Methods.

CHAPTER ONE

# *Introduction – Meaning and Scope*

**1·1. Origin and Development of Statistics.** Statistics, in a sense, is as old as the human society itself. Its origin can be traced to the old days when it was regarded as the 'science of State-craft' and was the by-product of the administrative activity of the State. The word 'Statistics' seems to have been derived from the Latin word '*status*' or the Italian word '*statista*' or the German word '*statistik*' each of which means a 'political state'. In ancient times, the government used to collect the information regarding the population and 'property or wealth' of the country – the former enabling the government to have an idea of the manpower of the country (to safeguard itself against external aggression, if any), and the latter providing it a basis for introducing news taxes and levies.

In India, an efficient system of collecting official and administrative statistics existed even more than 2,000 years ago, in particular, during the reign of Chandra Gupta Maurya ( 324 - 300 B.C.). From Kautilya's *Arthshastra* it is known that even before 300 B.C. a very good system of collecting 'Vital Statistics' and registration of births and deaths was in vogue. During Akbar's reign ( 1556 - 1605 A.D.), Raja Todarmal, the then land and revenue minister, maintained good records of land and agricultural statistics. In *Aina-e-Akbari* written by Abul Fazl (in 1596 - 97 ), one of the nine gems of Akbar, we find detailed accounts of the administrative and statistical surveys conducted during Akbar's reign.

In Germany, the systematic collection of official statistics originated towards the end of the 18th century when, in order to have an idea of the relative strength of different German States, information regarding population and output – industrial and agricultural – was collected. In England, statistics were the outcome of Napoleonic Wars. The Wars necessitated the systematic collection of numerical data to enable the government to assess the revenues and expenditure with greater precision and then to levy new taxes in order to meet the cost of war.

Seventeenth century saw the origin of the 'Vital Statistics.' Captain John Grant of London (1620 - 1674 ), known as the 'father' of Vital Statistics, was the first man to study the statistics of births and deaths. Computation of mortality tables and the calculation of expectation of life at different ages by a number of persons, *viz.*, Casper Newman, Sir William Petty (1623 - 1687 ), James Dodson, Dr. Price, to mention only a few, led to the idea of 'life insurance' and the first life insurance institution was founded in London in 1698.

The theoretical development of the so-called modern statistics came during the mid-seventeenth century with the introduction of 'Theory of Probability' and 'Theory of Games and Chance', the chief contributors being mathematicians and gamblers of France, Germany and England. The French mathematician Pascal (1623 - 1662 ), after lengthy correspondence with another French mathematician

P. Fermat (1601 - 1665 ) solved the famous 'Problem of Points' posed by the gambler Chevalier de - Mere. His study of the problem laid the foundation of the theory of probability which is the backbone of the modern theory of statistics. Pascal also investigated the properties of the co-efficients of binomial expansions and also invented mechanical computation machine. Other notable contributors in this field are : James Bernouli ( 1654 - 1705 ), who wrote the first treatise on the 'Theory of Probability'; De-Moivre (1667 - 1754 ) who also worked on probabilities and annuities and published his important work "The Doctrine of Chances" in 1718, Laplace (1749 - 1827 ) who published in 1782 his monumental work on the theory of probability, and Gauss (1777 - 1855 ), perhaps the most original of all writers on statistical subjects, who gave the principle of least squares and the normal law of errors. Later on, most of the prominent mathematicians of 18th, 19th and 20th centuries, *viz.*, Euler, Lagrange, Bayes, A. Markoff, Khintchin, Kolmogoroff, to mention only a few, added to the contributions in the field of probability.

Modern veterans in the development of the subject are Englishmen. Francis Galton (1822-1921), with his works on 'regression', pioneered the use of statistical methods in the field of Biometry. Karl Pearson (1857-1936), the founder of the greatest statistical laboratory in England (1911), is the pioneer in correlational analysis. His discovery of the 'chi square test', the first and the most important of modern tests of significance, won for Statistics a place as a science. In 1908 the discovery of Student's '*t*' distribution by W.S. Gosset who wrote under the pseudonym of 'Student' ushered in an era of exact sample tests (small samples).

Sir Ronald A. Fisher (1890 - 1962), known as the 'Father of Statistics', placed Statistics on a very sound footing by applying it to various diversified fields, such as genetics, biometry, education, agriculture, etc. Apart from enlarging the existing theory, he is the pioneer in introducing the concepts of 'Point Estimation' (efficiency, sufficiency, principle of maximum likelihood, etc.), 'Fiducial Inference' and 'Exact Sampling Distributions.' He also pioneered the study of 'Analysis of Variance' and 'Design of Experiments.' His contributions won for Statistics a very responsible position among sciences.

**1·2. Definition of Statistics.** Statistics has been defined differently by different authors from time to time. The reasons for a variety of definitions are primarily two. *First*, in modern times the field of utility of Statistics has widened considerably. In ancient times Statistics was confined only to the affairs of State but now it embraces almost every sphere of human activity. Hence a number of old definitions which were confined to a very narrow field of enquiry were replaced by new definitions which are much more comprehensive and exhaustive. *Secondly*, Statistics has been defined in two ways. Some writers define it as '*statistical data*', *i.e.*, numerical statement of facts, while others define it as '*statistical methods*', *i.e.*, complete body of the principles and techniques used in collecting and analysing such data. Some of the important definitions are given below.

## Statistics as 'Statistical Data'

Webster defines Statistics as "*classified facts representing the conditions of the people in a State ... especially those facts which can be stated in numbers or in any other tabular or classified arrangement.*" This definition, since it confines Statistics only to the data pertaining to State, is inadequate as the domain of Statistics is much wider.

Bowley defines Statistics as "*numerical statements of facts in any department of enquiry placed in relation to each other.*"

A more exhaustive definition is given by Prof. Horace Secrist as follows :

"*By Statistics we mean aggregates of facts affected to a marked extent by multiplicity of causes numerically expressed, enumerated or estimated according to reasonable standards of accuracy, collected in a systematic manner for a pre-determined purpose and placed in relation to each other.*"

## Statistics as Statistical Methods

Bowley himself defines Statistics in the following three different ways :

*(i)* Statistics may be called the science of counting.

*(ii)* Statistics may rightly be called the science of averages.

*(iii)* Statistics is the science of the measurement of social organism, regarded as a whole in all its manifestations.

But none of the above definitions is adequate. The *first* because statistics is not merely confined to the collection of data as other aspects like presentation, analysis and interpretation, etc., are also covered by it. The *second*, because averages are only a part of the statistical tools used in the analysis of the data, others being dispersion, skewness, kurtosis, correlation, regression, etc. The *third*, because it restricts the application of Statistics to sociology alone while in modern days Statistics is used in almost all sciences – social as well as physical.

According to Boddington, "*Statistics is the science of estimates and probabilities.*" This also is an inadequate definition since probabilities and estimates constitute only a part of the statistical methods.

Some other definitions are :

"*The science of Statistics is the method of judging collective, natural or social phenomenon from the results obtained from the analysis or enumeration or collection of estimates.*" – **King.**

"*Statistics is the science which deals with collection, classification and tabulation of numerical facts as the basis for explanation, description and comparison of phenomenon.*" – **Lovitt.**

Perhaps the best definition seems to be one given by Croxton and Cowden, according to whom Statistics may be defined as "*the science which deals with the collection, analysis and interpretation of numerical data.*"

**1·3. Importance and Scope of Statistics.** In modern times, Statistics is viewed not as a mere device for collecting numerical data but as a means of developing sound techniques for their handling and analysis and drawing valid inferences from them. As such it is not confined to the affairs of the State but is intruding constantly into various diversified spheres of life – social, economic and political. It is now finding wide applications in almost all sciences – social as well as physical – such as biology, psychology, education, economics, business management, etc. It is hardly possible to enumerate even a single department of human activity where statistics does not creep in. It has rather become indispensable in all phases of human endeavour.

**Statistics and Planning.** Statistics is indispensable to planning. In the modern age which is termed as 'the age of planning', almost all over the world, goernments, particularly of the budding economies, are resorting to planning for the economic development. In order that planning is successful, it must be based soundly on the correct analysis of complex statistical data.

**Statistics and Economics.** Statistical data and technique of statistical analysis have proved immensely useful in solving a variety of economic problems, such as wages, prices, analysis of time series and demand analysis. It has also facilitated the development of economic theory. Wide applications of mathematics and statistics in the study of economics have led to the development of new disciplines called Economic Statistics and Econometrics.

**Statistics and Business.** Statistics is an indispensable tool of production control also. Business executives are relying more and more on statistical techniques for studying the needs and the desires of the consumers and for many other purposes. The success of a businessman more or less depends upon the accuracy and precision of his statistical forecasting. Wrong expectations, which may be the result of faulty and inaccurate analysis of various causes affecting a particular phenomenon, might lead to his disaster. Suppose a businessman wants to manufacture readymade garments. Before starting with the production process he must have an overall idea as to 'how many garments are to be manufactured', 'how much raw material and labour is needed for that', and 'what is the quality, shape, colour, size, etc., of the garments to be manufactured'. Thus the formulation of a production plan in advance is a must which cannot be done without having quantitative facts about the details mentioned above. As such most of the large industrial and commercial enterprises are employing trained and efficient statisticians.

**Statistics and Industry.** In industry, Statistics is very widely used in 'Quality Control'. In production engineering, to find whether the product is conforming to specifications or not, statistical tools, *viz.*, inspection plans, control charts, etc., are of extreme importance. In inspection plans we have to resort to some kind of sampling – a very important aspect of Statistics.

**Statistics and Mathematics.** Statistics and mathematics are very intimately related. Recent advancements in statistical techniques are the outcome of wide applications of advanced mathematics. Main contributors to statistics, namely, Bernouli, Pascal, Laplace, De-Moivre, Gauss, R. A. Fisher, to mention only a few, were primarily talented and skilled mathematicians. Statistics may be regarded as that branch of mathematics which provided us with systematic methods of analysing a large number of related numerical facts. According to Connor, " Statistics is a branch of Applied Mathematics which specialises in data." Increasing role of mathematics in statistical analysis has resulted in a new branch of Statistics called Mathematical Statistics.

**Statistics and Biology, Astronomy and Medical Science.** The association between statistical methods and biological theories was first studied by Francis Galton in his work in 'Regression'. According to Prof. Karl Pearson, the whole 'theory of heredity' rests on statistical basis. He says, " *The whole problem of evolution is a problem of vital statistics, a problem of longevity, of fertility, of health, of disease and it is impossible for the Registrar General to discuss the national mortality without an enumeration of the population, a classification of deaths and knowledge of statistical theory.*"

In **astronomy**, the theory of Gaussian 'Normal Law of Errors' for the study of the movement of stars and planets is developed by using the 'Principle of Least Squares'.

In **medical science** also, the statistical tools for the collection, presentation and analysis of observed facts relating to the causes and incidence of diseases and the results obtained from the use of various drugs and medicines, are of great importance. Moreover, the efficacy of a manufacutured drug or injection or medicine is tested by using the 'tests of significance' – ($t$-test).

**Statistics and Psychology and Education.** In education and psychology, too, Statistics has found wide applications, *e.g.*, to determine the reliability and validity of a test, 'Factor Analysis', etc., so much so that a new subject called 'Psychometry' has come into existence.

**Statistics and War.** In war, the theory of 'Decision Functions' can be of great assistance to military and technical personnel to plan 'maximum destruction with minimum effort'.

Thus, we see that the science of Statistics is associated with almost all the sciences – social as well as physical. Bowley has rightly said, " *A knowledge of Statistics is like a knowledge of foreign language or of algebra; it may prove of use at any time under any circumstance.*"

**1·4. Limitations of Statistics.** Statistics, with its wide applications in almost every sphere of human activity; is not without limitations. The following are some of its important limitations :

*(i) Statistics is not suited to the study of qualitative phenomenon.* Statistics, being a science dealing with a set of numerical data, is applicable to the study of only those subjects of enquiry which are capable of quantitative measurement. As such, qualitative phenomena like honesty, poverty, culture, etc., which cannot be expressed numerically, are not capable of direct statistical analysis. However, statistical techniques may be applied indirectly by first reducing the qualitative expressions to precise quantitative terms. For example, the intelligence of a group of candidates can be studied on the basis of their scores in a certain test.

*(ii) Statistics does not study individuals.* Statistics deals with an aggregate of objects and does not give any specific recognition to the individual items of a series. Individual items, taken separately, do not constitute statistical data and are meaningless for any statistical enquiry. For example, the individual figures of agricultural production, industrial output or national income of any country for a particular year are meaningless unless, to facilitate comparison, similar figures of other countries or of the same country for different years are given. Hence, statistical analysis is suited to only those problems where group characteristics are to be studied.

*(iii) Statistical laws are not exact.* Unlike the laws of physical and natural sciences, statistical laws are only approximations and not exact. On the basis of statistical analysis we can talk only in terms of probability and chance and not in terms of certainty. Statistical conclusions are not universally true – they are true only on an average. For example, let us consider the statement :" It has been found that 20 % of a certain surgical operations by a particular doctor are successful." The statement does not imply that if the doctor is to operate on 5 persons on any day and four of the operations have proved fatal, the fifth must be a success. It may happen that fifth man also dies of the operation or it may also happen that of the five operations on any day, 2 or 3 or even more may be successful. By the statement we mean that as number of operations becomes larger and larger we should expect, on the average, 20 % operations to be successful.

*(iv) Statistics is liable to be misused.* Perhaps the most important limitation of Statistics is that it must be used by experts. As the saying goes, " Statistical methods are the most dangerous tools in the hands of the inexperts. Statistics is one of those sciences whose adepts must exercise the self-restraint of an artist." The use of statistical tools by inexperienced and untrained persons might lead to very fallacious conclusions. One of the greatest shortcomings of Statistics is that they do not bear on their face the label of their quality and as such can be moulded and manipulated in any manner to support one's way of argument and reasoning. As King says, *" Statistics are like clay of which one can make a god or devil as one pleases."* The requirement of experience and skill for judicious use of statistical methods restricts their use to experts only and limits the chances of the mass popularity of this useful and important science.

**1·5. Distrust of Statistics.** We often hear the following interesting comments on Statistics :

*(i)* 'An ounce of truth will produce tons of Statistics',

*(ii)* 'Statistics can prove anything',

*(iii)* 'Figures do not lie. Liars figure',

*(iv)* 'If figures say so it can't be otherwise',

*(v)* 'There are three type of lies – lies, demand lies, and Statistics – wicked in the order of their naming, and so on.

Some of the reasons for the existence of such divergent views regarding the nature and function of Statistics are as follows :

*(i)* Figures are innocent, easily believable and more convincing. The facts supported by figures are psychologically more appealing.

*(ii)* Figures put forward for arguments may be inaccurate or incomplete and thus might lead to wrong inferences.

*(iii)* Figures, though accurate, might be moulded and manipulated by selfish persons to conceal the truth and present a distorted picture of facts to the public to meet their selfish motives. When the skilled talkers, writers or politicians through their forceful writings and speeches or the business and commercial enterprises through advertisements in the press mislead the public or belie their expectations by quoting wrong statistical statements or manipulating statistsical data for personal motives, the public loses its faith and belief in the science of Statistics and starts condemning it. We cannot blame the layman for his distrust of Statistics, as he, unlike statistician, is not in a position to distinguish between valid and invalid conclusions from statistical statements and analysis.

It may be pointed out that Statistics neither proves anything nor disproves anything. It is only a tool which if rightly used may prove extremely useful and if misused, might be disastrous. According to Bowley, *"Statistics only furnishes a tool, necessary though imperfect, which is dangerous in the hands of those who do not know its use and its deficiencies."* It is not the subject of Statistics that is to be blamed but those people who twist the numerical data and misuse them either due to ignorance or deliberately for personal selfish motives. As King points out, *"Science of Statistics is the most useful servant but only of great value to those who understand its proper use."*

We discuss below a few interesting examples of misrepresentation of statistical data.

*(i)* A statistical report : "The number of accidents taking place in the middle of the road is much less than the number of accidents taking place on its side. Hence it is safer to walk in the middle of the road." This conclusion is obviously wrong since we are not given the proportion of the number of accidents to the number of persons walking in the two cases.

*(ii)* "The number of students taking up Mathematics Honours in a University has increased 5 times during the last 3 years. Thus, Mathematics is gaining popularity among the students of the university." Again, the conclusion is faulty since we are not given any such details about the other subjects and hence comparative study is not possible.

*(iii)* "99% of the people who drink alcohol die before attaining the age of 100 years. Hence drinking is harmful for longevity of life." This statement, too, is incorrect since nothing is mentioned about the number of persons who do not drink alcohol and die before attaining the age of 100 years.

Thus, statistical arguments based on incomplete data often lead to fallacious conclusions.

# FREQUENCY DISTRIBUTIONS AND MEASURES OF CENTRAL TENDENCY

2·1. **Frequency Distributions.** When observations, discrete or continuous, are available on a single characteristic of a large number of individuals, often it becomes necessary to condense the data as far as possible without losing any information of interest. Let us consider the marks in Statistics obtained by 250 candidates selected at random from among those appearing in a certain examination.

TABLE 1 : MARKS IN STATISTICS OF 250 CANDIDATES

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 32 | 47 | 41 | 51 | 41 | 30 | 39 | 18 | 48 | 53 |
| 54 | 32 | 31 | 46 | 15 | 37 | 32 | 56 | 42 | 48 |
| 38 | 26 | 50 | 40 | 38 | 42 | 35 | 22 | 62 | 51 |
| 44 | 21 | 45 | 31 | 37 | 41 | 44 | 18 | 37 | 47 |
| 68 | 41 | 30 | 52 | 52 | 60 | 42 | 38 | 38 | 34 |
| 41 | 53 | 48 | 21 | 28 | 49 | 42 | 36 | 41 | 29 |
| 30 | 33 | 37 | 35 | 29 | 37 | 38 | 40 | 32 | 49 |
| 43 | 32 | 24 | 38 | 38 | 22 | 41 | 50 | 17 | 46 |
| 46 | 50 | 26 | 15 | 23 | 42 | 25 | 52 | 38 | 46 |
| 41 | 38 | 40 | 37 | 40 | 48 | 45 | 30 | 28 | 31 |
| 40 | 33 | 42 | 36 | 51 | 42 | 56 | 44 | 35 | 38 |
| 31 | 51 | 45 | 41 | 50 | 53 | 50 | 32 | 45 | 48 |
| 40 | 43 | 40 | 34 | 34 | 44 | 38 | 58 | 49 | 28 |
| 40 | 45 | 19 | 24 | 34 | 47 | 37 | 33 | 37 | 36 |
| 36 | 32 | 61 | 30 | 44 | 43 | 50 | 31 | 38 | 45 |
| 46 | 40 | 32 | 34 | 44 | 54 | 35 | 39 | 31 | 48 |
| 48 | 50 | 43 | 55 | 43 | 39 | 41 | 48 | 53 | 34 |
| 32 | 31 | 42 | 34 | 34 | 32 | 33 | 24 | 43 | 39 |
| 40 | 50 | 27 | 47 | 34 | 44 | 34 | 33 | 47 | 42 |
| 17 | 42 | 57 | 35 | 38 | 17 | 33 | 46 | 36 | 23 |
| 48 | 50 | 31 | 58 | 33 | 44 | 26 | 29 | 31 | 37 |
| 47 | 55 | 57 | 37 | 41 | 54 | 42 | 45 | 47 | 43 |
| 37 | 52 | 47 | 46 | 44 | 50 | 44 | 38 | 42 | 19 |
| 52 | 45 | 23 | 41 | 47 | 33 | 42 | 24 | 48 | 39 |
| 48 | 44 | 60 | 38 | 38 | 44 | 38 | 43 | 40 | 48 |

This representation of the data does not furnish any useful information and is rather confusing to mind. A better way may be to express the figures in an ascending or descending order of magnitude, commonly termed as *array*. But this does not reduce the bulk of the data. A much better representation is given on the next page.

A bar ( | ) called *tally mark* is put against the number when it occurs. Having occurred four times, the fifth occurrence is represented by putting a cross tally (/) on the first four tallies. This technique facilitates the counting of the tally marks at the end.

The representation of the data as above is known as *frequency distribution.* Marks are called the *variable* ( $x$ ) and the *'number of students'* against the marks is known as the *frequency* ($f$) of the variable. The word *'frequency'* is derived from 'how frequently' a variable occurs. For example, in the above case the frequency of 31 is 10 as there are ten students getting 31 marks. This representation, though better than an array', does not condense the data much and it is quite cumbersome to go through this huge mass of data.

TABLE 2

| *Marks* | *No. of Students Tally Marks* | *Total Frequency* | *Marks* | *No. of Students Tally Marks* | *Total Frequency* |
|---|---|---|---|---|---|
| 15 | \|\| | = 2 | 40 | 卌 卌 \| | = 11 |
| 17 | \|\|\| | = 3 | 41 | 卌 卌 | = 10 |
| 18 | \|\| | = 2 | 42 | 卌 卌 \|\|\| | = 13 |
| 19 | \|\| | = 2 | 43 | 卌 \|\|\| | = 8 |
| 21 | \|\| | = 2 | 44 | 卌 卌 \|\| | = 12 |
| 22 | \|\| | = 2 | 45 | 卌 \|\| | = 7 |
| 23 | \|\|\| | = 3 | 46 | 卌 \|\| | = 7 |
| 24 | \|\|\|\| | = 4 | 47 | 卌 \|\|\| | = 8 |
| 25 | \| | = 1 | 48 | 卌 卌 \|\| | = 12 |
| 26 | \|\|\| | = 3 | 49 | \|\|\| | = 3 |
| 27 | \| | = 1 | 50 | 卌 卌 | = 10 |
| 28 | \|\|\| | = 3 | 51 | \|\|\|\| | = 4 |
| 29 | \|\| | = 2 | 52 | 卌 | = 5 |
| 30 | 卌 | = 5 | 53 | \|\|\|\| | = 4 |
| 31 | 卌 卌 | = 10 | 54 | \|\|\| | = 3 |
| 32 | 卌 卌 | = 10 | 55 | \|\| | = 2 |
| 33 | 卌 \|\|\| | = 8 | 56 | \|\| | = 2 |
| 34 | 卌 卌 \| | = 11 | 57 | \|\| | = 2 |
| 35 | 卌 | = 5 | 58 | \|\| | = 2 |
| 36 | 卌 | = 5 | 60 | \|\|\| | = 3 |
| 37 | 卌 卌 \|\| | = 12 | 61 | \| | = 1 |
| 38 | 卌 卌 卌 \|\| | = 17 | 62 | \| | = 1 |
| 39 | 卌 \| | = 6 | 68 | \| | = 1 |

If the identity of the individuals about whom a particular information is taken is not relevant, nor the order in which the observations arise, then the first real step of condensation is to divide the observed range of variable into a suitable number of *class-intervals* and to record the number of observations in each class. For example, in the above case, the data may be expressed as shown in Table 3.

Such a table showing the distribution of the frequencies in the different classes is called a *frequency table* and the manner in which the class frequencies are distributed over the class intervals is called the *grouped frequency distribution* of the variable.

**Remark.** The classes of the type 15—19, 20—24, 25—29 etc., in which both the upper and lower limits are included are called *'inclusive classes'*. For example the class 20—24, includes

TABLE 3 : FREQUENCY TABLE

| *Marks* ( $x$ ) | *No. of students.* ($f$) |
|---|---|
| 15 — 19 | 9 |
| 20 — 24 | 11 |
| 25 — 29 | 10 |
| 30 — 34 | 44 |
| 35 — 39 | 45 |
| 40 — 44 | 54 |
| 45 — 49 | 37 |
| 50 — 54 | 26 |
| 55 — 59 | 8 |
| 60 — 64 | 5 |
| 65 — 69 | 1 |
| Total | 250 |

all the values from 20 to 24, both inclusive and the classification is termed as *inclusive type classification.*

In spite of great importance of classification in statistical analysis, no hard and fast rules can be laid down for it. The following points may be kept in mind for classification :

*(i)* The classes should be clearly defined and should not lead to any ambiguity.

*(ii)* The classes should be exhaustive, *i.e.*, each of the given values should be included in one of the classes.

*(iii)* The classes should be mutually exclusive and non-overlapping.

*(iv)* The classes should be of equal width. The principle, however, cannot be rigidly followed. If the classes are of varying width, the different class frequencies will not be comparable. Comparable figures can be obtained by dividing the value of the frequencies by the corresponding widths of the class intervals. The ratios thus obtained are called *'frequency densities'*.

*(v)* Indeterminate classes, *e.g.*, the open-end classes, less than '$a$' or greater than '$b$' should be avoided as far as possible since they create difficulty in analysis and interpretation.

*(vi)* The number of classes should neither be too large nor too small. It should preferably lie between 5 and 15. However, the number of classes may be more than 15 depending upon the total frequency and the details required, but it is desirable that it is not less than 5 since in that case the classification may not reveal the essential characteristics of the population. The following formula due to Struges may be used to determine an approximate number $k$ of classes :

$$k = 1 + 3{\cdot}322 \log_{10} N,$$

where $N$ is the total frequency.

**The Magnitude of the Class Interval**

Having fixed the number of classes, divide the range ( the difference between the greatest and the smallest observation) by it and the nearest integer to this value gives the magnitude of the class interval. Broad class intervals ( *i.e.*, less number of classes) will yield only rough estimates while for high degree of accuracy small class intervals ( *i.e.*, large number of classes) are desirable.

**Class Limits**

The class limits should be chosen in such a way that the mid-value of the class interval and actual average of the observations in that class interval are as near to each other as possible. If this is not the case then the classification gives a distorted picture of the characteristics of the data. If possible, class limits should be located at the points which are multiple of 0, 2, 5, 10,... etc., so that the midpoints of the classes are the common figures, *viz.*, 0, 2, 5, 10..., etc., the figures capable of easy and simple analysis.

**2·1·1. Continuous Frequency Distribution.** If we deal with a continuous variable, it is not possible to arrange the data in the class intervals of above type. Let us consider the distribution of age in years. If class intervals are 15—19, 20—24 then the persons with ages between 19 and 20 years are not taken into consideration. In such a case we form the class intervals as shown below.

*Age in years*
Below 5
5 or more but less than 10
10 or more but less than 15
15 or more but less than 20
20 or more but less than 25
and so on.

Here all the persons with any fraction of age are included in one group or the other. For practical purpose we re-write the above classes as

0 — 5
5 — 10
10 — 15
15 — 20
20 — 25

This form of frequency distribution is known as *continuous frequency distribution.*

It should be clearly understood that in the above classes, the upper limits of each class are excluded from the respective classes. Such classes in which the upper limits are excluded from the respective classes and are included in the immediate next class are known as *'exclusive classes'* and the classification is termed as *'exclusive type classification'*.

**2·2. Graphic Representation of a Frequency Distribution.** It is often useful to represent a frequency distribution by means of a diagram which makes the unwieldy data intelligible and conveys to the eye the general run of the observations. Diagrammatic representation also facilitates the comparison of two or more frequency distributions. We consider below some important types of graphic representation.

**2·2·1. Histogram.** In drawing the histogram of a given continuous frequency distribution we first mark off along the $x$–axis all the class intervals on a suitable scale. On each class interval erect rectangles with heights proportional to the frequency of the corresponding class interval so that the area of the rectangle is proportional to the frequency of the class. If, however, the classes are of unequal width then the height of the rectangle wili be proportional to the ratio of the frequencies to the width of the classes. The diagram of continuous rectangles so obtained is called *histrogram.*

**Remarks.** 1. To draw the histogram for an ungrouped frequency distribution of a variable we shall have to assume that the frequency corresponding to the variate value $x$ is spread over the interval $x - h/2$ to $x + h/2$, where $h$ is the jump from one value to the next.

2. If the grouped frequency distribution is not continuous, first it is to be converted into continuous distribution and then the histrogam is drawn.

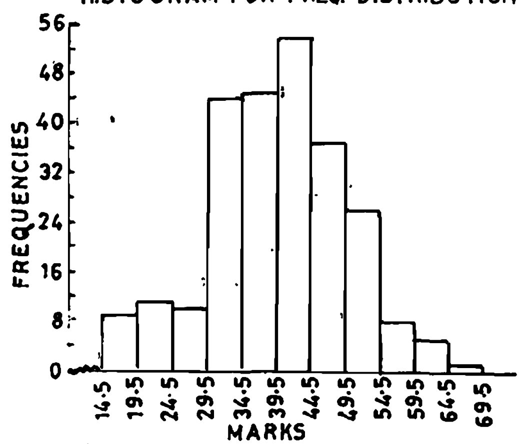
3. Although the height of each rectangle is proportional to the frequency of the corresponding class, the height of a fraction of the rectangle is not proportional to the frequency of the corresponding fraction of the class, so that histogram cannot be directly used to read frequency over a fraction of a class interval.

4. The histogram of the distribution of marks of 250 students in *Table* 3 (page 2·2) is obtained as follows.

Since the grouped frequency distribution is not continuous, we first convert it into a continuous distribution as follows:

| *Marks* | *No. of Students* |
|---|---|
| 14·5–19·5 | 9 |
| 19·5–24·5 | 11 |
| 24·5–29·5 | 10 |
| 29·5–34·5 | 44 |
| 34·5–39·5 | 45 |
| 39·5–44·5 | 54 |
| 44·5–49·5 | 37 |
| 49·5–54·5 | 26 |
| 54·5–59·5 | 8 |
| 59·5–64·5 | 5 |
| 64·5–69·5 | 1 |

HISTOGRAM FOR FREQ. DISTRIBUTION




**Remark.** The upper and lower class limits of the *new exclusive type classes* are known as *class boundaries.*

If $d$ is the gap between the upper limit of any class and the lower limit of the succeeding class, the class boundaries for any class are then given by :

$$\text{Upper class boundary} = \text{Upper class limit} + \frac{d}{2}$$

$$\text{Lower class boundary} = \text{Lower class limit} - \frac{d}{2}$$

2·2·2. **Frequency Polygon.** For an ungrouped distribution, the frequency polygon is obtained by plotting points with abscissa as the variate values and the ordinate as the corresponding frequencies and joining the plotted points by means of straight lines. For a grouped frequency distribution, the abscissa of points are mid-values of the class intervals. For equal class intervals the frequency polygon can be obtained by joining the middle points of the upper sides of the adjacent rectangles of the histogram by means of straight lines. If the class intervals are of small width the polygon can be approximated by a smooth curve. The frequency curve can be obtained by drawing a smooth freehand curve through the vertices of the frequency polygon.

**2·3. Averages or Measures of Central Tendency or Measures of Location.** According to Professor Bowley, averages are "statistical constants which enable us to comprehend in a single effort the significance of the whole." They give us an idea about the concentration of the values in the central part of the distribution. Plainly speaking, an average of a statistical series is the value of the variable which is representative of the entire distribution. The following are the five measures of central tendency that are in common use:

*(i) Arithmetic Mean or simply Mean , (ii) Median,*
*(iii) Mode, (iv) Geometric Mean, and (v) Harmonic Mean.*

**2·4. Requisites for an Ideal Measure of Central Tendency.** According to Professor Yule, the following are the characteristics to be satisfied by an ideal measure of central tendency :

*(i)* It should be rigidly defined.

*(ii)* It should be readily comprehensible and easy to calculate.

*(iii)* It should be based on all the observations.

*(iv)* It should be suitable for further mathematical treatment. By this we mean that if we are given the averages and sizes of a number of series, we should be able to calculate the average of the composite series obtained on combining the given series.

*(v)* It should be affected as little as possible by fluctuations of sampling.

In addition to the above criteria, we may add the following (which is not due to Prof.Yule) :

*(vi)* It should not be affected much by extreme values.

**2·5. Arithmetic Mean.** Arithmetic mean of a set of observastions is their sum divided by the number of observations, *e.g.*, the arithmetic mean $\bar{x}$ of $n$ observations $x_1$ , $x_2$ , ... , $x_n$ is given by

$$\bar{x} = \frac{1}{n}\left( x_1 + x_2 + \dots + x_n \right) = \frac{1}{n} \sum_{i=1}^{n} x_i$$

In case of frequency distribution $x_i \mid f_i$ , $i = 1, 2, ..., n$, where $f_i$ is the frequency of the variable $x_i$ ;

$$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \dots + f_n x_n}{f_1 + f_2 + \dots f_n} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{1}{N} \sum_{i=1}^{n} f_i x_i, \quad \left[ \sum_{i=1}^{n} f_i = N \right] \qquad ...(2·1)$$

In case of grouped or continuous frequency distribution, $x$ is taken as the mid-value of the corresponding class.

**Remark.** The symbol $\Sigma$ is the letter capital sigma of the Greek alphabet and is used in mathematics to denote the sum of values.

**Example 2·1.** *(a) Find the arithmetic mean of the following frequency distribution:*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| $f$ : | 5 | 9 | 12 | 17 | 14 | 10 | 6 |

*(b) Calculate the arithmetic mean of the marks from the following table :*

| | | | | | | |
|---|---|---|---|---|---|---|
| *Marks* : | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 | 50–60 |
| *No. of students* : | 12 | 18 | 27 | 20 | 17 | 6 |

**Solution.** *(a)*

| $x$ | $f$ | $fx$ |
|---|---|---|
| 1 | 5 | 5 |
| 2 | 9 | 18 |
| 3 | 12 | 36 |
| 4 | 17 | 68 |
| 5 | 14 | 70 |
| 6 | 10 | 60 |
| 7 | 6 | 42 |
| | 73 | 299 |

$$\therefore \qquad \bar{x} = \frac{1}{N} \Sigma fx = \frac{299}{73} = 4 \cdot 09$$

*(b)*

| *Marks* | *No. of students* ($f$) | *Mid - point* ($x$) | $fx$ |
|---|---|---|---|
| 0–10 | 12 | 5 | 60 |
| 10–20 | 18 | 15 | 270 |
| 20–30 | 27 | 25 | 675 |
| 30–40 | 20 | 35 | 700 |
| 40–50 | 17 | 45 | 765 |
| 50–60 | 6 | 55 | 330 |
| Total | 100 | | 2,800 |

Arithmetic mean or $\bar{x} = \frac{1}{N} \Sigma fx = \frac{1}{100} \times 2{,}800 = 28$

It may be noted that if the values of $x$ or (and) $f$ are large, the calculation of mean by formula *(2·1)* is quite time-consuming and tedious. The arithmetic is reduced to a great extent by taking the deviations of the given values from any arbitrary point '$A$', as explained below.

Let $d_i = x_i - A$, then $f_i d_i = f_i (x_i - A) = f_i x_i - A f_i$

Summing both sides over $i$ from 1 to $n$, we get

$$\sum_{i=1}^{n} f_i d_i = \sum_{i=1}^{n} f_i x_i - A \sum_{i=1}^{n} f_i = \sum_{i=1}^{n} f_i x_i - A \, . N \cdot$$

$$\Rightarrow \qquad \frac{1}{N}\sum_{i=1}^{n} f_i d_i = \frac{1}{N}\sum_{i=1}^{n} f_i x_i - A = \bar{x} - A\,,$$

where $\bar{x}$ is the arithmetic mean of the distribution.

$$\therefore \qquad \bar{x} = A + \frac{1}{N}\sum_{i=1}^{n} f_i d_i \qquad \ldots(2{\cdot}2)$$

This formula is much more convenient to apply than formula (2·1).

Any number can serve the purpose of arbitrary point '$A$' but, usually, the value of $x$ corresponding to the middle part of the distribution will be much more convenient.

In case of grouped or continuous frequency distribution, the arithmetic is reduced to a still greater extent by taking

$$d_i = \frac{x_i - A}{h}\,,$$

where $A$ is an arbitrary point and $h$ is the common magnitude of class interval. In this case, we have

$$h\,d_i = x_i - A\,,$$

and proceeding exactly similarly as above, we get

$$\bar{x} = A + \frac{h}{N}\sum_{i=1}^{n} f_i d_i \qquad \ldots(2{\cdot}3)$$

**Example 2·2.** *Calculate the mean for the following frequency distribution.*

| | | | | | | |
|---|---|---|---|---|---|---|
| *Class-interval* : | *0–8* | *8–16* | *16–24* | *24–32* | *32–40* | *40–48* |
| *Frequency* : | *8* | *7* | *16* | *24* | *15* | *7* |

**Solution.**

| *Class–interval* | *mid–value* $(x)$ | *Frequency* $(f)$ | $d = (x-A)/h$ | $fd$ |
|---|---|---|---|---|
| 0–8 | 4 | 8 | –3 | –24 |
| 8–16 | 12 | 7 | –2 | –14 |
| 16–24 | 20 | 16 | –1 | –16 |
| 24–32 | 28 | 24 | 0 | 0 |
| 32–40 | 36 | 15 | 1 | 15 |
| 40–48 | 44 | 7 | 2 | 14 |
| | | 77 | | –25 |

Here we take $A = 28$ and $h = 8$.

$$\therefore \qquad \bar{x} = A + \frac{h\,\Sigma f d}{N} = 28 + \frac{8 \times (-25)}{77} = 28 - \frac{200}{77} = 25{\cdot}404$$

## 2·5·1. Properties of Arithmetic Mean

**Property 1.** *Algebraic sum of the deviations of a set of values from their arithmetic mean is zero. If* $x_i\,|\,f_i$, $i = 1, 2, \ldots, n$ *is the frequency distribution, then*

$$\sum_{i=1}^{n} f_i \, (x_i - \bar{x}) = 0, \; \bar{x} \textit{ being the mean of distribution.}$$

**Proof.** $\sum_i f_i (x_i - \bar{x}) = \sum_i f_i x_i - \bar{x} \sum f_i = \sum_i f_i x_i - \bar{x} \, . N$

Also $\bar{x} = \dfrac{\sum_i f_i x_i}{N} \Rightarrow \sum_i f_i x_i = N \bar{x}$

Hence $\sum_{i=1}^{n} f_i (x_i - \bar{x}) = N . \bar{x} - \bar{x} . N = 0$

**Property 2.** *The sum of the squares of the deviations of a set of values is minimum when taken about mean.*

**Proof.** For the frequency distribution $x_i \,|\, f_i$, $i = 1, 2, ..., n$, let

$$Z = \sum_{i=1}^{n} f_i (x_i - A)^2,$$

be the sum of the squares of the deviations of given values from any arbitrary point '$A$'. We have to prove that $Z$ is minimum when $A = \bar{x}$.

Applying the principle of maxima and minima from differential calculus, $Z$ will be minimum for variations in $A$ if

$$\frac{\partial Z}{\partial A} = 0 \text{ and } \frac{\partial^2 Z}{\partial A^2} > 0$$

Now $\dfrac{\partial Z}{\partial A} = - 2 \sum_i f_i (x_i - A) = 0 \Rightarrow \sum_i f_i (x_i - A) = 0$

$$\Rightarrow \quad \sum f_i x_i - A \sum f_i = 0 \quad or \quad A = \frac{\sum f_i x_i}{N} = \bar{x}$$

Again $\dfrac{\partial^2 Z}{\partial A^2} = - 2 \sum_i f_i (-1) = 2 \sum_i f_i = 2N > 0$

Hence $Z$ is minimum at the point $A = \bar{x}$. This establishes the result.

**Property 3.** *(Mean of the composite series). If $\bar{x}_i$, ($i = 1, 2, ..., k$) are the means of k-component series of sizes $n_i$, ($i = 1, 2, ..., k$) respectively, then the mean $\bar{x}$ of the composite series obtained on combining the component series is given by the formula:*

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + \ldots + n_k \bar{x}_k}{n_1 + n_2 + \ldots + n_k} = \sum_i n_i \bar{x}_i \Big/ \sum_i n_i \qquad ...(2\text{·}4)$$

**Proof.** Let $x_{11}, x_{12}, ... x_{1n_1}$ be $n_1$ members of the first series; $x_{21}, x_{22}, ...., x_{2n_2}$ be $n_2$ members of the second series, $x_{k1}, x_{k2}, ..., x_{kn_k}$ be $n_k$ members of the $k$th series. Then, by def.,

$$\left.\begin{aligned} \bar{x}_1 &= \frac{1}{n_1}(x_{11} + x_{12} + \dots + x_{1n_1}) \\ \bar{x}_2 &= \frac{1}{n_2}(x_{21} + x_{22} + \dots + x_{2n_2}) \\ &\vdots \qquad \vdots \qquad \vdots \\ \bar{x}_k &= \frac{1}{n_k}(x_{k1} + x_{k2} + \dots + x_{kn_k}) \end{aligned}\right\} \qquad \dots(*)$$

The mean $\bar{x}$ of composite series of size $n_1 + n_2 + \dots + n_k$ is given by

$$\bar{x} = \frac{(x_{11} + x_{12} + \dots + x_{1n_1}) + (x_{21} + x_{22} + \dots + x_{2n_2}) + \dots + (x_{k1} + x_{k2} + \dots + x_{kn_k})}{n_1 + n_2 + \dots + n_k}$$

$$= \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + \dots + n_k \bar{x}_k}{n_1 + n_2 + \dots n_k}, \qquad \text{[From (*)]}$$

Thus, $\bar{x} = \sum\limits_i n_i \bar{x}_i \;/\; (\sum\limits_i n_i)$

**Example 2·3.** *The average salary of male employees in a firm was Rs.520 and that of females was Rs.420. The mean salary of all the employees was Rs.500. Find the percentage of male and female employees.*

**Solution.** Let $n_1$ and $n_2$ denote respectively the number of male and female employees in the concern and $\bar{x}_1$ and $\bar{x}_2$ denote respectively their average salary (in rupees). Let $\bar{x}$ denote the avarage salary of all the workers in the firm.

We are given that :

$$\bar{x}_1 = 520, \quad \bar{x}_2 = 420 \quad \text{and} \quad \bar{x} = 500$$

Also we know

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2}$$

$$\Rightarrow \quad 500\,(n_1 + n_2) = 520\,n_1 + 420\,n_2$$

$$\Rightarrow \quad (520 - 500)\,n_1 = (500 - 420)\,n_2$$

$$\Rightarrow \quad 20\,n_1 = 80\,n_2$$

$$\Rightarrow \quad \frac{n_1}{n_2} = \frac{4}{1}$$

Hence the percentage of male employees in the firm

$$= \frac{4}{4+1} \times 100 = 80$$

and percentage of female employees in the firm

$$= \frac{1}{4+1} \times 100 = 20$$

## 2·5·2. Merits and Demerits of Arithmetic Mean

**Merits.** *(i)* It is rigidly defined.

*(ii)* It is easy to understand and easy to calculate.

*(iii)* It is based upon all the observations.

*(iv)* It is amenable to algebraic treatment. The mean of the composite series in terms of the means and sizes of the component series is given by

$$\bar{x} = \sum_{i=1}^{k} n_i \bar{x}_i \Big/ \left( \sum_{i}^{k} n_i \right)$$

*(v)* Of all the averages, arithmetic mean is affected least by fluctuations of sampling. This property is sometimes described by saying that arithmetic mean is a *stable* average.

Thus, we see that arithmetic mean satisfies all the properties laid down by Prof. Yule for an ideal average.

**Demerits.** *(i)* It cannot be determined by inspection nor it can be located graphically.

*(ii)* Arithmetic mean cannot be used if we are dealing with qualitative characteristics which cannot be measured quantitively; such as, intelligence, honesty, beauty, etc. In such cases median (discussed later) is the only average to be used.

*(iii)* Arithmetic mean cannot be obtained if a single observation is missing or lost or is illegible unless we drop it out and compute the arithmetic mean of the remaining values.

*(iv)* Arithmetic mean is affected very much by extreme values. In case of extreme items, arithmetic mean gives a distorted picture of the distribution and no longer remains representative of the distribution.

*(v)* Arithmetic mean may lead to wrong conclusions if the details of the data from which it is computed are not given. Let us consider the following marks obtained by two students *A* and *B* in three tests, *viz.*, terminal test, half-yearly examination and annual examination respectively.

| *Marks in* : → | *I Test* | *II Test* | *III Test* | *Average marks* |
|---|---|---|---|---|
| *A* | 50% | 60% | 70% | 60% |
| *B* | 70% | 60% | 50% | 60% |

Thus average marks obtained by each of the two students at the end of the year are 60%. If we are given the average marks alone we conclude that the level of intelligence of both the students at the end of the year is same. This is a fallacious conclusion since we find from the data that student *A* has improved consistently while student *B* has deteriorated consistently.

*(vi)* Arithmetic mean cannot be calculated if the extreme class is open, *e.g.*, below 10 or above 90. Morever, even if a single observation is missing mean cannot be calculated.

*(vii)* In extremely asymmertrical (skewed) distribution, usually arithmetic mean is not a suitable measure of location.

**2·5·3. Weighted Mean.** In calculating arithmetic mean we suppose that all the items in the distribution have equal importance. But in practice this may not be so. If some items in a distribution are more important than others, then this point must be borne in mind, in order that average computed is representative of the distribution. In such cases, proper weightage is to be given to various items — the weights attached to each item being proportional to the importance of the item in the distribution. For example, if we want to have an idea of the change in cost of living of a certain group of people, then the simple mean of the prices of the commodities consumed by them will not do, since all the commodities are not equally important, *e.g.,* wheat, rice and pulses are more important than cigarettes, tea, confectionery, etc.

Let $w_i$ be the weight attached to the item $x_i$, $i = 1, 2, \ldots, n$. Then we define :

Weighted arithmetic mean or weighted mean $= \sum\limits_i w_i x_i \,/\, \sum\limits_i w_i$ ... (2·5)

It may be observed that the formula for weighted mean is the same as the formula for simple mean with $f_i$, $(i = 1, 2, \ldots, n)$, *the frequencies replaced by* $w_i$, $(i = 1, 2, \ldots, n)$, *the weights.*

Weighted mean gives the result equal to the simple mean if the weights assigned to each of the variate values are equal. It results in higher value than the simple mean if smaller weights are given to smaller items and larger weights to larger items. If the weights attached to larger items are smaller and those attached to smaller items are larger, then the weighted mean results in smaller value than the simple mean.

**Example 2·4.** *Find the simple and weighted arithmetic mean of the first $n$ natural numbers, the weights being the corresponding numbers.*

**Solution.** The first *natural* numbers are $1, 2, 3, \ldots, n$.

| $X$ | $w$ | $wX$ |
|---|---|---|
| 1 | 1 | $1^2$ |
| 2 | 2 | $2^2$ |
| 3 | 3 | $3^2$ |
| ⋮ | ⋮ | ⋮ |
| $n$ | $n$ | $n^2$ |

We know that

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

$$1^2 + 2^2 + 3^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

Simple A.M. is

$$\overline{X} = \frac{\sum X}{n} = \frac{1 + 2 + 3 + \ldots + n}{n} = \frac{n+1}{2}$$

Weighted A.M. is

$$\overline{X}_w = \frac{\sum wX}{\sum w} = \frac{1^2 + 2^2 + \ldots + n^2}{1 + 2 + \ldots + n}$$

$$= \frac{n(n+1)(2n+1)}{6} \cdot \frac{2}{n(n+1)}$$

**2·6. Median.** Median of a distribution is the value of the variable which divides it into two equal parts. It is the value which exceeds and is exceeded by the same number of observations, *i.e.*, it is the value such that the number of observations above it is equal to the number of observations below it. The median is thus a *positional* average.

In case of ungrouped data, if the number of observations is odd then median is the middle value after the values have been arranged in ascending or descending order of magnitude. In case of even number of observations, there are two middle terms and median is obtained by taking the arithmetic mean of the middle terms. For example, the median of the values 25, 20, 15, 35, 18, *i.e.*, 15, 18, 20, 25, 35 is 20 and the median of 8, 20, 50, 25, 15, 30, *i.e.*, of 8, 15, 20, 25, 30, 50 is $\frac{1}{2}(20+25) = 22{\cdot}5$.

**Remark.** In case of even number of observations, in fact any value lying between the two middle values can be taken as median but conventionally we take it to be the mean of the middle terms.

In case of discrete frequency distribution median is obtained by considering the cumulative frequencies. The steps for calculating median are given below:

*(i)* Find $N/2$, where $N = \sum_i f_i$.

*(ii)* See the (less than) cumulative frequency (*c.f.*) just greater than $N/2$.

*(iii)* The corresponding value of $x$ is median.

**Example 2·5.** *Obtain the median for the following frequency distribution:*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| $f$ : | 8 | 10 | 11 | 16 | 20 | 25 | 15 | 9 | 6 |

**Solution.**

| $x$ | $f$ | $c.f.$ |
|---|---|---|
| 1 | 8 | 8 |
| 2 | 10 | 18 |
| 3 | 11 | 29 |
| 4 | 16 | 45 |
| 5 | 20 | 65 |
| 6 | 25 | 90 |
| 7 | 15 | 105 |
| 8 | 9 | 114 |
| 9 | 6 | 120 |
| | 120 | |

Hence $N = 120 \quad \Rightarrow \quad N/2 = 60$

Cumulative frequency (*c.f.*) just greater than $N/2$, is 65 and the value of $x$ corresponding to 65 is 5. Therefore, median is 5.

In the case of continuous frequency distribution, the class corresponding to the *c.f.* just greater than $N/2$ is called the *median class* and the value of median is obtained by the following formula :

$$\text{Median} = l + \frac{h}{f}\left(\frac{N}{2} - c\right) \qquad ...(2\cdot6)$$

where $l$ is the lower limit of the median class,
$f$ is the frequency of the median class,
$h$ is the magnitude of the median class,
'$c$' is the *c.f.* of the class *preceding* the median class,
and $N = \Sigma f$.

**Derivation of the Median Formula (2·6).** Let us consider the following continuous frequency distribution, $(x_1 < x_2 < \ldots < x_{n+1})$ :

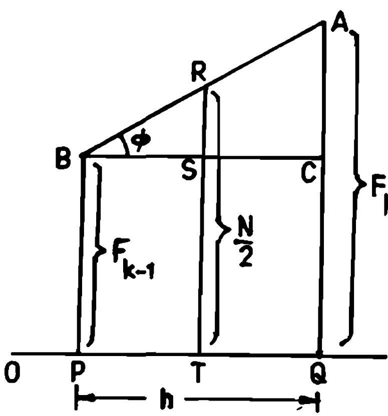
Class interval : $x_1 - x_2$, $x_2 - x_3$, ........ $x_k - x_{k+1}$, ........ $x_n - x_{n+1}$
Frequency : $f_1$ $f_2$ ........ $f_k$ ........ $f_n$

The cumulaltive frequency distribution is given by :

Class interval : $x_1 - x_2$, $x_2 - x_3$, ........ $x_k - x_{k+1}$, ........ $x_n - x_{n+1}$
Frequency : $F_1$ $F_2$ ........ $F_k$ ........ $F_n$

where $F_i = f_1 + f_2 + \ldots\ldots + f_i$. The class $x_k - x_{k+1}$ is the median class if and only if $F_{k-1} < N/2 \le F_k$.

Now, if we assume that the variate values are uniformly distributed over the median-class which implies that the ogive is a straight line in the median-class, then we get from the Fig. 1,




$$\tan\phi = \frac{RS}{BS} = \frac{AC}{BC}$$

*i.e.*
$$\frac{RT - TS}{BS} = \frac{AQ - CQ}{BC}$$

*or*
$$\frac{RT - BP}{BS} = \frac{AQ - BP}{PQ}$$

*or*
$$\frac{N/2 - F_{k-1}}{BS} = \frac{F_k - F_{k-1}}{PQ} = \frac{f_k}{h}$$

where $f_k$ is the frequency and $h$ the magnitude of the median class.

$$\therefore \qquad BS = \frac{h}{f_k}\left(\frac{N}{2} - F_{k-1}\right)$$

Hence

$$\text{Median} = OT = OP + PT = OP + BS$$
$$= l + \frac{h}{f_k}\left(\frac{N}{2} - F_{k-1}\right)$$

which is the required formula.

**Remark.** The median formula (2·6) can be used only for continuous classes without any gaps, *i.e.*, for '*exclusive type*' classification. If we are given a frequency distribution in which classes are of *'inclusive type'* with gaps, then it must be converted into a continuous *'exclusive type'* frequency distribution without any gaps before applying (2·6). This will affect the value of $l$ in (2·6). As an illustration see Example 2·7.

**Example 2·6.** *Find the median wage of the following distribution :*

| | | | | | |
|---|---|---|---|---|---|
| *Wages (in Rs.) :* | *20—30* | *30—40* | *40—50* | *50—60* | *60—70* |
| *No. of labourers :* | *3* | *5* | *20* | *10* | *5* |

**[Gorakhpur Univ. B. Sc. 1989]**

**Solution.**

| *Wages (in Rs.)* | *No. of labourers* | *c.f.* |
|---|---|---|
| 20—30 | 3 | 3 |
| 30—40 | 5 | 8 |
| 40—50 | 20 | 28 |
| 50—60 | 10 | 38 |
| 60—70 | 5 | 43 |

Here $N/2 = 43/2 = 21·5$ . Cumulative frequency just greater than 21·5 is 28 and the corresponding class is 40–50. Thus median class is 40–50. Hence using (2·6), we get

$$\text{Median} = 40 + \frac{10}{20}(21·5 - 8) = 40 + 6·75 = 46·75$$

Thus median wage is Rs. 46·75.

**Example 2·7.** *In a factory employing 3,000 persons, 5 per cent earn less than Rs. 3 per hour, 580 earn from Rs. 3·01 to Rs. 4·50 per hour, 30 percent earn from Rs. 4·51 to Rs. 6·00 per hour, 500 earn from Rs. 6·01 to Rs. 7·50 per hour, 20 percent earn from Rs. 7·51 to Rs. 9·00 per hour, and the rest earn Rs. 9·01 or more per hour. What is the median wage?* **[Utkal Univ. B.Sc.1992]**

**Solution.** The given information can be expressed in tabular form as follows.

CALCULATIONS FOR MEDIAN WAGE

| *Earnings (in Rs.)* | *Percentage of workers* | *No. of workers (f)* | *Less than c.f.* | *Class boundaries* |
|---|---|---|---|---|
| less than 3 | 5% | $\frac{5}{100} \times 3000 = 150$ | 150 | Below 3·005 |
| 3·01—4·50 | — | 580 | 730 | 3·005—4·505 |
| 4·51—6·00 | 30 % | $\frac{30}{100} \times 3000 = 900$ | 1630 | 4·505—6·005 |
| 6·01—7·50 | — | 500 | 2130 | 6·005—7·505 |
| 7·51—9·00 | 20 % | $\frac{20}{100} \times 3000 = 600$ | 2730 | 7·505—9·005 |
| 9·01 and above | | $3000 - 2730 = 270$ | $3000 = N$ | 9·005 and above |

$N/2 = 1500$. The $c.f.$ just greater than 1500 is 1630. The corresponding class 4·51–6·00, whose class boundaries are 4·505–6·005, is the median class. Using the median formula, we get:

$$\text{Median} = l + \frac{h}{f}\left(\frac{N}{2} - C\right) = 4{\cdot}505 + \frac{1{\cdot}5}{900}(1500 - 730)$$
$$= 4{\cdot}505 + 1{\cdot}283 \approx 5{\cdot}79$$

Hence median wage is Rs. 5·79.

**Example 2·8.** *An incomplete frequency distribution is given as follows:*

| *Variable* | *Frequency* | *Variable* | *Frequency* |
|---|---|---|---|
| *10—20* | *12* | *50—60* | *?* |
| *20—30* | *30* | *60—70* | *25* |
| *30—40* | *?* | *70—80* | *18* |
| *40—50* | *65* | *Total* | *229* |

*Given that the median value is 46, determine the missing frequencies using the median formula.* **[Delhi Univ. B. Sc., Oct. 1992]**

**Solution.** Let the frequency of the class 30—40 be $f_1$ and that of 50—60 be $f_2$.

Then $f_1 + f_2 = 229 - (12 + 30 + 65 + 25 + 18) = 79$.

Since median is given to be 46, the class 40—50 is the median class.

Hence using median formula (2·6), we get

$$46 = 40 + \frac{114{\cdot}5 - (12 + 30 + f_1)}{65} \times 10$$

$$46 - 40 = \frac{72.5 - f_1}{65} \times 10 \quad \text{or} \quad 6 = \frac{72{\cdot}5 - f_1}{6{\cdot}5}$$

$$f_1 = 72{\cdot}5 - 39 = 33{\cdot}5 \approx 34$$

[Since frequency is never fractional]

$\therefore \quad f_2 = 79 - 34 = 45$ [Since $f_1 + f_2 = 79$]

**2·6·1. Merits and Demerits of Median**

**Merits.** *(i)* It is rigidly defined.

*(ii)* It is easily understood and is easy to calculate. In some cases it can be located merely by inspection.

*(iii)* It is not at all affected by extreme values.

*(iv)* It can be calculated for distributions with open-end classes.

**Demerits.** *(i)* In case of even number of observations median cannot be determined exactly. We merely estimate it by taking the mean of two middle terms.

*(ii)* It is not based on all the observations. For example, the median of 10, 25, 50, 60 and 65 is 50. We can replace the observations 10 and 25 by any two values which are smaller than 50 and the observations 60 and 65 by any two values greater than 50 without affecting the value of median. This property is sometimes described

by saying that median is *insensitive*.

*(iii)* It is not amenable to algebraic treatment.

*(iv)* As compared with mean, it is affected much by fluctuations of sampling.

**Uses.** *(i)* Median is the only average to be used while dealing with qualitative data which cannot be measured quantitatively but still can be arranged in ascending or descending order of magnitude, *e.g.*, to find the average intelligence or average honesty among a group of people.

*(ii)* It is to be used for determining the typical value in problems concerning wages, distribution of wealth, etc.

**2·7. Mode.** Let us cosider the following statements :

*(i)* The average height of an Indian (male) is 5′–6″.

*(ii)* The average size of the shoes sold in a shop is 7.

*(iii)* An average student in a hostel spends Rs.150 p.m.

In all the above cases, the average referred to is mode. Mode is the value which occurs most frequently in a set of observastions and around which the other items of the set cluster densely. In other words, mode is the value of the variable which is predominant in the series. Thus in the case of discrete frequency distribution mode is the value of $x$ corresponding to maximum frequency. For example, in the following frequency distribution :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $f$ : | 4 | 9 | 16 | 25 | 22 | 15 | 7 | 3 |

the value of $x$ corresponding to the maximum frequency, *viz.*, 25 is 4. Hence mode is 4.

But in any one (or more) of the following cases :

*(i)* if the maximum frequency is repeated,

*(ii)* if the maximum frequency occurs in the very beginning or at the end of the distribution, and

*(iii)* if there are irregularities in the distribution,

the value of mode is determined by the *method of grouping*, which is illustrated below by an example.

**Example 2·9.** *Find the mode of the following frequency distribution :*

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Size* ($x$) : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| *Frequency* ($f$) : | 3 | 8 | 15 | 23 | 35 | 40 | 32 | 28 | 20 | 45 | 14 | 6 |

**Solution.** Here we see that the distribution is not regular since the frequencies are increasing steadily up to 40 and then decrease but the frequency 45 after 20 does not seem to be consistent with the distribution. Here we cannot say that since maximum frequency is 45, mode is 10. Here we shall locate mode by the method of grouping as explained below :

The frequencies in column *(i)* are the original frequencies. Column *(ii)* is obtained by combining the frequencies two by two. If we leave the first frequency and combine the remaining frequencies two by two we get column *(iii)*. Combining

| Size $(x)$ | Frequency (i) | (ii) | (iii) | (iv) | (v) | (vi) |
|---|---|---|---|---|---|---|
| 1 | 3 | 11 | | 26 | | |
| 2 | 8 | | 23 | | 46 | |
| 3 | 15 | 38 | | | | 73 |
| 4 | 23 | | 58 | **98** | | |
| 5 | 35 | **75** | | | **107** | |
| 6 | 40 | | **72** | | | **100** |
| 7 | 32 | 60 | | 80 | | |
| 8 | 28 | | 48 | | 93 | |
| 9 | 20 | 65 | | | | 79 |
| 10 | **45** | | 59 | 65 | | |
| 11 | 14 | 20 | | | | |
| 12 | 6 | | | | | |

the frequencies two by two after leaving the first two frequencies results in a repetition of column *(ii)*. Hence, we proceed to combine the frequencies three by three, thus getting column *(iv)*. The combination of frequencies three by three after leaving the first frequency results in column *(v)* and after leaving the first two frequencies results in column *(vi)*.

The maximum frequency in each column is given in black type. To find mode we form the following table :

ANALYSIS TABLE

| *Column Number* *(1)* | *Maximum Frequency* *(2)* | *Value or combination of values of x giving max. frequency in (2)* *(3)* |
|---|---|---|
| *(i)* | 45 | 10 |
| *(ii)* | 75 | 5, 6 |
| *(iii)* | 72 | 6, 7 |
| *(iv)* | 98 | 4, 5, 6, |
| *(v)* | 107 | 5, 6, 7 |
| *(vi)* | 100 | 6, 7, 8 |

On examining the values in column (3) above, we find that the value 6 is repeated the maximum number of times and hence the value of mode is 6 and not 10 which is an irregular item.

In case of continuous frequency distribution, mode is given by the formula :

$$Mode = l + \frac{h(f_1 - f_0)}{(f_1 - f_0) - (f_2 - f_1)} = l + \frac{h(f_1 - f_0)}{2f_1 - f_0 - f_2} \qquad ...(2\cdot7)$$

where $l$ is the lower limit, $h$ the magnitude and $f_1$ the frequency of the modal class, $f_0$ and $f_2$ are the frequencies of the classes preceding and succeeding the modal class respectively.

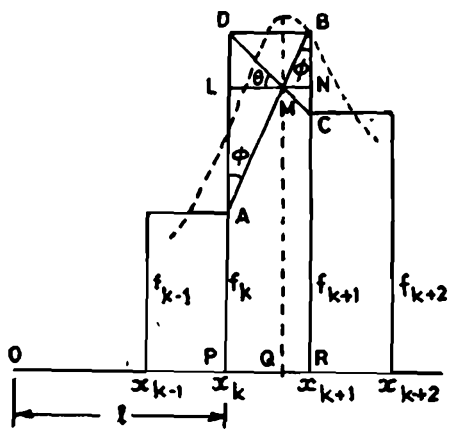
**Derivation of the Mode Formula (2·7).** Let us consider the continuous frequency distribution :

Class : $x_1 — x_2$, $x_2 — x_3$, ... ... , $x_k — x_{k+1}$, ... ... , $x_n — x_{n+1}$

Frequency : $f_1$ $f_2$ ... ... $f_k$ ... ... $f_n$.

If $f_k$ is the maximum of all the frequencies, then the modal class is $(x_k — x_{k+1})$.

Let us further consider a portion of the histogram, namely, the rectangles erected on the modal class and the two adjacent classes. The mode is the value of $x$ for which the frequency curve has a maxima. Let the modal point be Q.




From the figure, we have

$$\tan \theta = \frac{LD}{LM} = \frac{NC}{MN}$$

and
$$\tan \phi = \frac{LM}{AL} = \frac{MN}{NB}$$

$\therefore$
$$\frac{LM}{MN} = \frac{LD}{NC} = \frac{AL}{NB} = \frac{AL + LD}{NB + NC} = \frac{AD}{BC}$$

*i.e.,*
$$\frac{LM}{LN - LM} = \frac{PD - AP}{BR - CR}$$

or
$$\frac{LM}{h - LM} = \frac{f_k - f_{k-1}}{f_k - f_{k+1}},$$ where '$h$' is the magnitude of the modal class. Thus solving for $LM$, we get

$$LM = \frac{h(f_k - f_{k-1})}{(f_k - f_{k+1}) + (f_k - f_{k-1})} = \frac{h(f_k - f_{k-1})}{2f_k - f_{k-1} - f_{k+1}}$$

Hence Mode $= OQ = OP + PQ = OP + LM$

$$= l + \frac{h(f_k - f_{k-1})}{2f_k - f_{k-1} - f_{k+1}}$$

**Example 2·10.** *Find the mode for the following distribution :*

| *Class - interval :* | *0–10* | *10–20* | *20–30* | *30–40* | *40–50* | *50–60* | *60–70* | *70–80* |
|---|---|---|---|---|---|---|---|---|
| *Frequency :* | *5* | *8* | *7* | *12* | *28* | *20* | *10* | *10* |

**Solution.** Here maximum frequency is 28. Thus the class 40–50 is the modal class. Using (2·7), the value of mode is given by

$$\text{Mode} = 40 + \frac{10(28-12)}{(2 \times 28 - 12 - 20)} = 40 + 6{\cdot}666 = 46{\cdot}67 \quad \text{(approx.)}$$

**Example 2·11.** *The Median and Mode of the following wage distribution are known to be Rs. 33·50 and Rs. 34 respectively. Find the values of* $f_3$, $f_4$ *and* $f_5$.

| *Wages : (in Rs.)* | *0–10* | *10–20* | *20–30* | *30–40* | *40–50* |
|---|---|---|---|---|---|
| *Frequency :* | *4* | *16* | $f_3$ | $f_4$ | $f_5$ |

| *Wages :* | *50–60* | *60–70* | *Total* |
|---|---|---|---|
| *Frequency :* | *6* | *4* | *230* |

**[Gujarat Univ. B.Sc., 1991]**

**Solution.**

CALCULATIONS FOR MODE AND MEDIAN

| *Wages (in Rs.)* | *Frequency (f)* | *Less than c.f.* |
|---|---|---|
| 0—10 | 4 | 4 |
| 10—20 | 16 | 20 |
| 20—30 | $f_3$ | $20 + f_3$ |
| 30—40 | $f_4$ | $20 + f_3 + f_4$ |
| 40—50 | $f_5$ | $20 + f_3 + f_4 + f_5$ |
| 50—60 | 6 | $26 + f_3 + f_4 + f_5$ |
| 60—70 | 4 | $30 + f_3 + f_4 + f_5$ |
| *Total* | $230 = 30 + f_3 + f_4 + f_5$ | |

From the above table, we get

$$\Sigma f = 30 + f_3 + f_4 + f_5 = 230$$

$$\Rightarrow \quad f_3 + f_4 + f_5 = 230 - 30 = 200 \qquad ...(i)$$

Since median is 33·5, which lies in the class 30–40, 30–40 is the median class. Using the median formula, we get

$$Md = l + \frac{h}{f}\left(\frac{N}{2} - C\right)$$

$$\Rightarrow \qquad 33{\cdot}5 = 30 + \frac{10}{f_4}\left[115 - (20 + f_3)\right]$$

$$\Rightarrow \qquad \frac{33{\cdot}5 - 30}{10} = \frac{95 - f_3}{f_4}$$

$$\Rightarrow \qquad 0{\cdot}35 f_4 = 95 - f_3 \quad \Rightarrow \quad f_3 = 95 - 0{\cdot}35 f_4 \qquad \text{...}(ii)$$

Mode being 34, the modal class is also 30–40. Using mode formula we get :

$$34 = 30 + \frac{10\,(f_4 - f_3)}{2f_4 - f_3 - f_5}$$

$$\Rightarrow \qquad \frac{34 - 30}{10} = \frac{f_4 + 0{\cdot}35\, f_4 - 95}{2f_4 - (200 - f_4)} \qquad \text{[Using } (i) \text{ and } (ii)\text{]}$$

$$\Rightarrow \qquad 0{\cdot}4 = \frac{1{\cdot}35 f_4 - 95}{3 f_4 - 200}$$

$$\Rightarrow \qquad 1{\cdot}2 f_4 - 80 = 1{\cdot}35 f_4 - 95$$

$$\Rightarrow \qquad f_4 = \frac{95 - 80}{1{\cdot}35 - 1{\cdot}20} = \frac{15}{0{\cdot}15} = 100 \qquad \text{...}(iii)$$

Substituting in *(ii)* we get :

$$f_3 = 95 - 0{\cdot}35 \times 100 = 60$$

Substituting the values of $f_3$ and $f_4$ in *(i)* we get :

$$f_5 = 200 - f_3 - f_4 = 200 - 60 - 100 = 40$$

Hence $\qquad f_3 = 60, f_4 = 100$ and $f_5 = 40.$

**Remarks. 1.** In case of irregularities in the distribution, or the maximum frequency being repeated or the maximum frequency occurring in the very beginning or at the end of the distribution, the modal class is determined by the method of grouping and the mode is obtained by using (2·7).

Sometimes mode is estimated from the mean and the median. For a symmetrical distribution, mean, median and mode coincide. If the distribution is *moderately asymmetrical,* the mean, median and mode obey the following empirical relationship ( due to Karl Pearson) :

$$\text{Mean} - \text{Median} = \tfrac{1}{3}\,(\text{Mean} - \text{Mode})$$

$$\Rightarrow \qquad \text{Mode} = 3\ \text{Median} - 2\ \text{Mean} \qquad \text{...}(2{\cdot}8)$$

**2.** If the method of grouping gives the modal class which does not correspond to the maximum frequency, *i.e.,* the frequency of modal class is not the maximum frequency, then in some situations we may get, $2f_k - f_{k-1} - f_{k+1} = 0$. In such cases, the value of mode can be obtained by the formula :

$$\text{Mode} = l + \frac{h\,(f_k - f_{k-1})}{|f_k - f_{k-1}| + |f_k - f_{k+1}|}$$

**2·7·1. Merits and Demerits of Mode**

**Merits.** *(i)* Mode is readily comprehensible and easy to calculate. Like median, mode can be located in some cases merely by inspection.

*(ii)* Mode is not at all affected by extreme values.

*(iii)* Mode can be conveniently located even if the frequency distribution has class-intervals of unequal magnitude provided the modal class and the classes preceding and succeeding it are of the same magnitude. Open–end classes also do not pose any problem in the location of mode.

**Demerits.** *(i)* Mode is ill-defined. It is not always possible to find a clearly defined mode. In some cases, we may come across distributions with two modes. Such distributions are called *bi-modal*. If a distribution has more than two modes, it is said to be *multimodal*.

*(ii)* It is not based upon all the observations.

*(iii)* It is not capable of further mathematical treatment.

*(iv)* As compared with mean, mode is affected to a greater extent by fluctuations of sampling.

**Uses.** Mode is the average to be used to find the ideal size, *e.g.*, in business forecasting, in the manufacture of ready-made garments, shoes, etc.

**2·8. Geometric Mean.** Geometric mean of a set of $n$ observations is the $n$th root of their product. Thus the geometric mean $G$, of $n$ observations $x_i$, $i = 1, 2, ..., n$ is

$$G = (x_1 \,.\, x_2 \,. \dots .\, x_n)^{1/n} \qquad \text{...(2·9)}$$

The computation is facilitated by the use of logarithms. Taking logarithm of both sides, we get

$$\log G = \frac{1}{n}(\log x_1 + \log x_2 + \dots + \log x_n) = \frac{1}{n}\sum_{i=1}^{n} \log x_i$$

$$\therefore \qquad G = \text{Antilog}\left[\frac{1}{n}\sum_{i=1}^{n} \log x_i\right] \qquad \text{...(2·9}a\text{)}$$

In case of frequency distribution $x_i \mid f_i$, ($i = 1, 2, ..., n$) geometric mean, $G$ is given by

$$G = \left[x_1^{f_1} \,.\, x_2^{f_2} \dots .\, x_n^{f_n}\right]^{\frac{1}{N}}, \text{ where } N = \sum_{i=1}^{n} f_i \qquad \text{...(2·10)}$$

Taking logarithms of both sides, we get

$$\log G = \frac{1}{N}(f_1 \log x_1 + f_2 \log x_2 + \dots + f_n \log x_n)$$

$$= \frac{1}{N}\sum_{i=1}^{n} f_i \log x_i \qquad \text{...(2·10}a\text{)}$$

Thus we see that logarithm of $G$ is the arithmetic mean of the logarithms of the given values. From (2·10$a$), we get

$$G = \text{Antilog}\left(\frac{1}{N}\sum_{i=1}^{n} f_i \log x_i\right) \qquad \ldots(2\cdot10b)$$

In the case of grouped or continuous frequency distribution, $x$ is taken to be the value corresponding to the mid-point of the class-intervals.

### 2·8·1. Merits and Demerits of Geometric Mean

**Merits.** *(i)* It is rigidly defined.

*(ii)* It is based upon all the observations.

*(iii)* It is suitable for further mathematical treatment. If $n_1$ and $n_2$ are the sizes, $G_1$ and $G_2$ the geometric means of two series respectively, the geometric mean $G$, of the combined series is given by

$$\log G = \frac{n_1 \log G_1 + n_2 \log G_2}{n_1 + n_2} \qquad \ldots(2\cdot11)$$

**Proof.** Let $x_{1i}$ $(i = 1, 2, \ldots, n_1)$ and $x_{2j}$ $(j = 1, 2, \ldots, n_2)$ be $n_1$ and $n_2$ items of two series respectively. Then by def.,

$$G_1 = (x_{11} . x_{12} \ldots x_{1n1})^{1/n_1} \quad \Rightarrow \quad \log G_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} \log x_{1i}$$

$$G_2 = (x_{21} . x_{22} \ldots x_{2n2})^{1/n_2} \quad \Rightarrow \quad \log G_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} \log x_{2j}$$

The geometric mean $G$ of the combined series is given by

$$G = (x_{11} . x_{12} \ldots x_{1n1} . x_{21} . x_{22} \ldots x_{2n2})^{1/(n_1+n_2)}$$

$$\therefore \qquad \log G = \frac{1}{n_1+n_2}\left[\sum_{i=1}^{n_1} \log x_{1i} + \sum_{j=1}^{n_2} \log x_{2j}\right]$$

$$= \frac{1}{n_1+n_2}\left[n_1 \log G_1 + n_2 \log G_2\right]$$

The result can be easily generalised to more than two series.

*(iv)* It is not affected much by fluctuations of sampling.

*(v)* It gives comparatively more weight to small items.

**Demerits.** *(i)* Because of its abstract mathematical character, geometric mean is not easy to understand and to calculate for a non-mathematics person.

*(ii)* If any one of the observations is zero, geometric mean becomes zero and if any one of the observations is negative, geometric mean becomes imaginary regardless of the magnitude of the other items.

**Uses.** Geometric mean is used –

*(i)* To find the rate of population growth and the rate of interest.

*(ii)* In the construction of index numbers.

**Example 2·12.** *Show that in finding the arithmetic mean of a set of readings on thermometer it does not matter whether we measure temperature in Centigrade or Fahrenheit, but that in finding the geometric mean it does matter which scale we use.* **[Patna Univ. B.Sc., 1991]**

**Solution.** Let $C_1, C_2, ..., C_n$ be the $n$ readings on the Centigrade thermometer. Then their arithmetic mean $\overline{C}$ is given by :

$$\overline{C} = \frac{1}{n}(C_1 + C_2 + \ldots + C_n)$$

If $F$ and $C$ be the readings in Fahrenheit and Centigrade respectively then we have the relation :

$$\frac{F-32}{180} = \frac{C}{100} \quad \Rightarrow \quad F = 32 + \frac{9}{5}C.$$

Thus the Fahrenheit equivalents of $C_1, C_2, ..., C_n$ are

$$32 + \frac{9}{5}C_1,\ 32 + \frac{9}{5}C_2,\ \ldots,\ 32 + \frac{9}{5}C_n,$$

respectively.

Hence the arithmetic mean of the readings in Fahrenheit is

$$\begin{aligned}\overline{F} &= \frac{1}{n}\left\{\left(32 + \frac{9}{5}C_1\right) + \left(32 + \frac{9}{5}C_2\right) + \ldots + \left(32 + \frac{9}{5}C_n\right)\right\}\\ &= \frac{1}{n}\left\{32n + \frac{9}{5}(C_1 + C_2 + \ldots + C_n)\right\}\\ &= 32 + \frac{9}{5}\left(\frac{C_1 + C_2 + \ldots + C_n}{n}\right)\\ &= 32 + \frac{9}{5}\overline{C}.\end{aligned}$$

which is the Fahrenheit equivalent of $\overline{C}$.

Hence in finding the arithmetic mean of a set of $n$ readings on a thermometer, it is immaterial whether we measure temperature in Centigrade or Fahrenheit.

Geometric mean $G$, of $n$ readings in Centigrade is

$$G = (C_1 . C_2 \ldots C_n)^{1/n}$$

Geometric mean $G_1$, (say), of Fahrenheit equivalents of $C_1, C_2, ..., C_n$ is

$$G_1 = \left\{\left(32 + \frac{9}{5}C_1\right)\left(32 + \frac{9}{5}C_2\right)\ldots\left(32 + \frac{9}{5}C_n\right)\right\}^{1/n}$$

which is not equal to Fahrenheit equivalent of $G$, *viz.*,

$$\left\{\frac{9}{5}(C_1 . C_2 \ \ldots \ C_n)^{1/n} + 32\right\}$$

Hence in finding the geometric mean of the $n$ readings on a thermometer, the scale (Centrigrade or Fahrenheit) is important.

**2·9. Harmonic Mean.** Harmonic mean of a number of observations is the reciprocal of the arithmetic mean of the reciprocals of the given values. Thus, harmonic mean $H$, of $n$ observations $x_i$, $i = 1, 2, ..., n$ is

$$H = \frac{1}{\frac{1}{n} \sum_{i=1}^{n} (1/x_i)} \qquad ...(2\cdot12)$$

In case of frequency distribution $x_i \mid f_i$, $(i = 1, 2, ..., n)$,

$$H = \frac{1}{\frac{1}{N} \sum_{i=1}^{n} (f_i/x_i)}, \quad \left[N = \sum_{i=1}^{n} f_i\right] \qquad ...(2\cdot12a)$$

**2·9·1. Merits and Demerits of Harmonic Mean**

**Merits.** Harmonic mean is rigidly defined, based upon all the observations and is suitable for further mathematical treatment. Like geometric mean, it is not affected much by fluctuations of sampling. It gives greater importance to small items and is useful only when small items have to be given a greater weightage.

**Demerits.** Harmonic mean is not easily understood and is difficult to compute.

**Example 2·13.** *A cyclist pedals from his house to his college at a speed of 10 m.p.h. and back from the college to his house at 15 m.p.h. Find the average speed.*

**Solution.** Let the distance from the house to the college be $x$ miles. In going from house to college, the distance ($x$ miles) is covered in $\frac{x}{10}$ hours, while in coming from college to house, the distance is covered in $\frac{x}{15}$ hours. Thus a total distance of $2x$ miles is covered in $\left(\frac{x}{10} + \frac{x}{15}\right)$ hours.

$$\text{Hence average speed} = \frac{\text{Total distance travelled}}{\text{Total time taken}} = \frac{2x}{\left(\frac{x}{10} + \frac{x}{15}\right)}$$

$$= \frac{2}{\left(\frac{1}{10} + \frac{1}{15}\right)} = 12 \text{ } m.p.h.$$

**Remark. 1.** In this case the average speed is given by the harmonic mean of 10 and 15 and not by the airthmetic mean.

Rather, we have the following general result :

If equal distances are covered (travelled) per unit of time with speeds equal to $V_1, V_2, ..., V_n$, say, then the average speed is given by the harmonic mean of $V_1, V_2, ..., V_n$, *i.e.,*

$$\text{Average speed} = \frac{n}{\left(\frac{1}{V_1} + \frac{1}{V_2} + ... + \frac{1}{V_n}\right)} = \frac{n}{\Sigma\left(\frac{1}{V}\right)}$$

Proof is left as an exercise to the reader.

**Hint.** Speed $= \dfrac{\text{Distance}}{\text{Time}} \Rightarrow$ Time $= \dfrac{\text{Distance}}{\text{Speed}}$

Average Speed $= \dfrac{\text{Total distance travelled}}{\text{Total time taken}}$

**2. Weighted Harmonic Mean.** Instead of fixed (constant) distance being travelled with varying speed, let us now suppose that different distances, say, $S_1, S_2, \ldots, S_n$, are travelled with different speeds, say, $V_1, V_2, \ldots, V_n$ respectively. In that case, the average speed is given by the weighted harmonic mean of the speeds, the weights being the corresponding distances travelled, *i.e.*,

$$\text{Average speed} = \frac{S_1 + S_2 + \ldots + S_n}{\left(\dfrac{S_1}{V_1} + \dfrac{S_2}{V_2} + \ldots + \dfrac{S_n}{V_n}\right)} = \frac{\Sigma S}{\Sigma \left(\dfrac{S}{V}\right)}$$

**Example 2·14.** *You can take a trip which entails travelling 900 km. by train at an average speed of 60 km. per hour, 3000 km. by boat at an average of 25 km. p.h., 400 km. by plane at 350 km. per hour and finally 15 km. by taxi at 25 km. per hour. What is your average speed for the entire distance ?*

**Solution.** Since different distances are covered with varying speeds, the required average speed for the entire distance is given by the weighted harmonic mean of the speeds (in km.p.h.), the weights being the corresponding distances covered (in kms.).

COMPUTATION OF WEIGHTED H. M.

| *Speed (km. / hr.)* $X$ | *Distance (in km.)* $W$ | $W/X$ |
|---|---|---|
| 60 | 900 | 15·00 |
| 25 | 3000 | 120·00 |
| 350 | 400 | 1·43 |
| 25 | 15 | 0·60 |
| Total | $\Sigma W = 4315$ | $\Sigma (W/X) = 137·03$ |

Average speed

$$= \frac{\Sigma W}{\Sigma (W/X)} = \frac{4315}{137·03} = 31·489 \text{ km.p.h.}$$

**2·10. Selection of an Average.** From the preceding discussion it is evident that no single average is suitable for all practical purposes. Each one of the average has its own merits and demerits and thus its own particular field of importance and utility. We cannot use the averages indiscriminately. A judicious selection of the average depending on the nature of the data and the purpose of the enquiry is essential for sound statistical analysis. Since arithmetic mean satisfies all the properties of an ideal average as laid down by Prof. Yule, is familiar to a layman and further has wide applications in statistical theory at large, it may be regarded as the best of all the averages.

**2·11. Partition Values.** These are the values which divide the series into a number of equal parts.

The three points which divide the series into four equal parts are called *quartiles*. The first, second and third points are known as the first, second and third quartiles respectively. The first quartile, $Q_1$, is the value which exceed 25% of the observations and is exceeded by 75% of the observations. The second quartile, $Q_2$, coincides with median. The third quartile, $Q_3$, is the point which has 75% observations before it and 25% observations after it.

The nine points which divide the series into ten equal parts are called *deciles* whereas *percentiles* are the ninety-nine points which divide the series into hundred equal parts. For example, $D_7$, the seventh decile, has 70% observations before it and $P_{47}$, the forty-seventh percentile, is the point which exceed 47% of the observations. The methods of computing the partition values are the same as those of locating the median in the case of both discrete and continuous distributions.

**Example 2·15.** *Eight coins were tossed together and the number of heads resulting was noted. The operation was repeated 256 times and the frequencies (f) that were obtained for different values of x, the number of heads, are shown in the following table. Calculate median, quartiles, 4th decile and 27th precentile.*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $f$ : | 1 | 9 | 26 | 59 | 72 | 52 | 29 | 7 | 1 |

**Solution.**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $f$ : | 1 | 9 | 26 | 59 | 72 | 52 | 29 | 7 | 1 |
| $c.f.$ : | 1 | 10 | 36 | 95 | 167 | 219 | 248 | 255 | 256 |

Median : Here $N/2 = 256/2 = 128$. Cumulative frequency $(c.f.)$ just greater than 128 is 167. Thus, median $= 4$.

$Q_1$ : Here $N/4 = 64$. $c.f.$ just greater than 64 is 95. Hence, $Q_1 = 3$.

$Q_3$ : Here $3N/4 = 192$ and $c.f.$ just greater than 192 is 219. Thus $Q_3 = 5$.

$D_4$ : $\dfrac{4N}{10} = 4 \times 25{\cdot}6 = 102{\cdot}4$ and $c.f.$ just greater than $102{\cdot}4$ is 167. Hence $D_4 = 4$.

$P_{27}$ : $\dfrac{27N}{100} = 27 \times 2{\cdot}56 = 69{\cdot}12$ and $c.f.$ just greater than $69{\cdot}12$ is 95. Hence $P_{27} = 3$.

**2·11·1. Graphical Location of the Partition Values.** The partition values, *viz.*, quartiles, deciles and percentiles, can be conveniently located with the help of a curve called the 'cumulative frequency curve' or 'Ogive'. The procedure is illustrated below.

First form the cumulative frequency table. Take the class intervals (or the variate values) along the $x$-axis and plot the corresponding cumulative frequencies along the $y$-axis against the *upper limit* of the class interval (or against the variate value in the case of discrete frequency distribution). The curve obtained on joining the points so obtained by means of free hand drawing is called the *cumulative frequency curve* or *ogive*. The graphical location of partition values from this curve is explained below by means of an example.

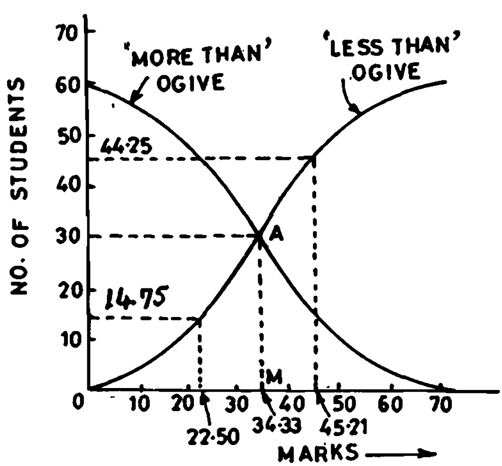
**Example 2·16.** *Draw the cumulative frequency curve for the following distribution showing the number of marks of 59 students in Statistics.*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Marks-group : | 0—10 | 10—20 | 20—30 | 30—40 | 40—50 | 50—60 | 60—70 |
| No. of Students : | 4 | 8 | 11 | 15 | 12 | 6 | 3 |

**Solution.**

| *Marks-group* | *No. of Students* | *Less than c.f.* | *More than c.f.* |
|---|---|---|---|
| 0—10 | 4 | 4 | 59 |
| 10—20 | 8 | 12 | 55 |
| 20—30 | 11 | 23 | 47 |
| 30—40 | 15 | 38 | 36 |
| 40—50 | 12 | 50 | 21 |
| 50—60 | 6 | 56 | 9 |
| 60—70 | 3 | 59 | 3 |

Taking the marks–group along $x$-axis and *c.f.* along $y$-axis, we plot the cumulative frequencies, *viz.*, 4, 12, 23, ..., 59 against the upper limits of the corresponding classes, *viz.*, 10, 20, ..., 70 respectively. The smooth curve obtained on joining these points is called *ogive* or more particularly *'less than' ogive.*




If we plot the 'more than' cumulative frequencies, *viz.*, 59, 55,..., 3 against the lower limits of the corresponding classes, *viz.*, 0, 10, ...; 60 and join the points by a smooth curve, we get cumulative frequency curve which is also known as *ogive* or more particularly *'more than' ogive.*

To locate graphically the value of median, mark a point corresponding to $N/2$ along $y$-axis. At this point draw a line parallel to $x$-axis meeting the ogive at the point '$A$' (say). From '$A$' draw a line prependicular to $x$-axis meeting it in '$M$' (say). Then abscissa of '$M$' gives the value of median.

To locate the values of $Q_1$ (or $Q_3$), we mark the points along $y$-axis corresponding to $N/4$ (or $3N/4$) and proceed exactly similarly.

In the above example, we get from ogive

Median= 34·33, $Q_1$ = 22·50, and $Q_3$ = 45·21.

**Remarks. 1.** The median can also be located as follows :

From the point of intersection of 'less than' ogive and 'more than' ogive, draw perpendicular to $OX$. The abscissa of the point so obtained gives median.

2. Other partition values, *viz.*, deciles and percentiles, can be similarly located from 'ogive'.

## EXERCISE

1. *(a)* What are grouped and ungrouped frequency distributions? What are their uses? What are the considerations that one has to bear in mind while forming the frequency distribution?

*(b)* Explain the method of constructing Histogram and Frequency Polygon. Which, out of these two, is better representative of frequencies of *(i)* a particular group, and *(ii)* whole group.

2. What are the principles governing the choice of :

*(i)* Number of class intervals,

*(ii)* The length of the class interval,

*(iii)* The mid-point of the class interval.

3. Write short notes on :

*(i)* Frequency distribution,

*(ii)* Histogram, frequency polygon and frequency curve,

*(iii)* Ogive.

4. *(a)* What are the properties of a good average? Examine these properties with reference to the Arithmetic Mean, the Geometric Mean and the Harmonic Mean, and give an example of situations in which each of them can be the appropriate measure for the average.

*(b)* Compare mean, median and mode as measures of location of a distribution.

*(c)* The mean is the most common measure of central tendency of the data. It satisfies almost all the requirements of a good average. The median is also an average, but it does not statisfy all the requirements of a good average. However, it carries certain merits and hence is useful in particular fields. Critically examine both the averages.

*(d)* Describe the different measures of central tendency of a frequency distribution, mentioning their merits and demerits.

5. Define *(i)* arithmetic mean, *(ii)* geometric mean and *(iii)* harmonic mean of grouped and ungrouped data. Compare and contrast the merits and demerits of them. Show that the geometric mean is capable of further mathematical treatment.

6. *(a)* When is an average a meaningful statistics? What are the requisites of a statisfactory average? In this light compare the relative merits and demerits of three well-known averages.

*(b)* What are the chief measures of central tendency? Discuss their merits.

7. Show that *(i)* Sum of deviations about arithmetic mean is zero.

*(ii)* Sum of absolute deviations about median is least.

*(iii)* Sum of the squares of deviations about arithmetic mean is least.

8. The following numbers give the weights of 55 students of a class. Prepare a suitable frequency table.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 42 | 74 | 40 | 60 | 82 | 115 | 41 | 61 | 75 | 83 | 63 |
| 53 | 110 | 76 | 84 | 50 | 67 | 65 | 78 | 77 | 56 | 95 |
| 68 | 69 | 104 | 80 | 79 | 79 | 54 | 73 | 59 | 81 | 100 |
| 66 | 49 | 77 | 90 | 84 | 76 | 42 | 64 | 69 | 70 | 80 |
| 72 | 50 | 79 | 52 | 103 | 96 | 51 | 86 | 78 | 94 | 71 |

*(i)* Draw the histogram and frequency polygon of the above data.

*(ii)* For the above weights, prepare a cumulative frequency table and draw the less than ogive.

9. *(a)* What are the points to be borne in mind in the formation of frequency table?

Choosing appropriate class-intervals, form a frequency table for the following data:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 10·2 | 0·5 | 5·2 | 6·1 | 3·1 | 6·7 | 8·9 | 7·2 | 8·9 |
| 5·4 | 3·6 | 9·2 | 6·1 | 7·3 | 2·0 | 1·3 | 6·4 | 8·0 |
| 4·3 | 4·7 | 12·4 | 8·6 | 13·1 | 3·2 | 9·5 | 7·6 | 4·0 |
| 5·1 | 8·1 | 1·1 | 11·5 | 3·1 | 6·8 | 7·0 | 8·2 | 2·0 |
| 3·1 | 6·5 | 11·2 | 12·0 | 5·1 | 10·9 | 11·2 | 8·5 | 2·3 |
| 3·4 | 5·2 | 10·7 | 4·9 | 6·2 | | | | |

*(b)* What are the considerations one has to bear in mind while forming a frequency distribution?

A sample consists of 34 observations recorded correct to the nearest integer, ranging in value from 201 to 337. If it is decided to use seven classes of width 20 integers and to begin the first class at 199·5, find the class limits and class marks of the seven classes.

*(c)* The class marks in a frequency table (of whole numbers) are given to be 5, 10, 15, 20, 25, 30, 35, 40, 45 and 50. Find out the following :

*(i)* the true classes.

*(ii)* the true class limits.

*(iii)* the true upper class limits.

**10.** *(a)* The following table shows the distribution of the number of students per teacher in 750 colleges :–

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Students : | 1 | 4 | 7 | 10 | 13 | 16 | 19 | 22 | 25 | 28 |
| Frequency : | 7 | 46 | 165 | 195 | 189 | 89 | 28 | 19 | 9 | 3 |

Draw the histogram for the data and superimpose on it the frequency polygon.

*(b)* Draw the histogram and frequency curve for the following data.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Monthly wages in Rs. | 10–13 | 13–15 | 15–17 | 17–19 | 19–21 | 21–23 | 23–25 |
| No. of workers | 6 | 53 | 85 | 56 | 21 | 16 | 8 |

*(c)* Draw a histogram for the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| Age (in years) : | 2–5 | 5–11 | 11–12 | 12–14 | 14–15 | 15–16 |
| No. of boys : | 6 | 6 | 2 | 5 | 1 | 3 |

**11.** *(a)* Three people $A, B, C$ were given the job of finding the average of 5000 numbers. Each one did his own simplification. $A$'s method : Divide the sets into sets of 1000 each, calculate the average in each set and then calculate the average of these averages. $B$'s method : Divide the set into 2,000 and 3,000 numbers, take average in each set and then take the average of the averages. $C$'s method :500 numbers were unities. He averaged all other numbers and then added one. Are these methods correct?

**Ans.** Correct, nôt correct, not correct.

*(b)* The total sale (in '000 rupees) of a particular item in a shop, on 10 consecutive days, is reported by a clerk as, 35·00, 29·60, 38·00, 30·00, 40·00, 41·00, 42·00, 45·00, 3·60, 3·80. Calculate the average. Later it was found that there was a number 10·00 in the machine and the reports of 4th to 8th days were 10·00 more than the true values and in the last 2 days he put a decimal in the wrong place thus for example 3·60 was really 36·0. Calculate the true mean value.

**Ans.** 30·8, 32·46.

**12.** *(a)* Given below is the distribution of 140 candidates obtaining marks $X$ or higher in a certain examination (all marks are given in whole numbers) :

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $X$ : | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
| $c.f.$ : | 140 | 133 | 118 | 100 | 75 | 45 | 25 | 9 | 2 | 0 |

Calculate the mean, median and mode of the distribution.

**Hint.**

| *Class* | *Frequency (f)* | *Class boundaries* | *Mid value* | *c.f. (less than)* |
|---|---|---|---|---|
| 10—19 | 140 – 133 = 7 | 9·5—19·5 | 14·5 | 7 |
| 20—29 | 133 – 118 = 15 | 19·5—29·5 | 24·5 | 22 |
| 30—39 | 118 – 100 = 18 | 29·5—39·5 | 34·5 | 40 |
| 40—49 | 100 – 75 = 25 | 39·5—49·5 | 44·5 | 65 |
| 50—59 | 75 – 45 = 30 | 49·5—59·5 | 54·5 | 95 |
| 60—69 | 45 – 25 = 20 | 59·5—69·5 | 64·5 | 115 |
| 70—79 | 25 – 9 = 16 | 69·5—79·5 | 74·5 | 131 |
| 80—89 | 9 – 2 = 7 | 79·5—89·5 | 84·5 | 138 |
| 90—99 | 2 – 0 = 2 | 89·5—99·5 | 94·5 | 140 |

$$\text{Mean} = 54 \cdot 5 + \frac{10 \times (-53)}{140} = 50 \cdot 714$$

$$\text{Median} = 49 \cdot 5 + \frac{10}{30}\left(\frac{140}{2} - 65\right) = 51 \cdot 167$$

*(b)* The four parts of a distribution are as follows :

| *Part* | *Frequency* | *Mean* |
|---|---|---|
| 1 | 50 | 61 |
| 2 | 100 | 70 |
| 3 | 120 | 80 |
| 4 | 30 | 83 |

Find the mean of the distribution. **(Madurai Univ. B.Sc., 1988)**

**13.** *(a)* Define a 'weighted mean'. If several sets of observations are combined into a single set, show that the mean of the combined set is the weighted mean of several sets.

*(b)* The weighted geometric mean of three numbers 229, 275 and 125 is 203 The weights for the first and second numbers are 2 and 4 respectively. Find the weight of third. Ans. 3.

**14.** Define the weighted arithmetic mean of a set of numbers. Show that it is unaffected if all weights are multiplied by some common factor.

The following table shows some data collected for the regions of a country:

| *Region* | *Number of inhabitants (million)* | *Percentage of literates* | *Average annual income per person (Rs.)* |
|---|---|---|---|
| *A* | 10 | 52 | 850 |
| *B* | 5 | 68 | 620 |
| *C* | 18 | 39 | 730 |

Obtain the overall figures for the three regions taken together. Prove the formulae you use. **[Calcutta Univ. B.A.(Hons.), 1991]**

**15.** Draw the Ogives and hence estimate the median.

| *Class* | 0–9 | 10–19 | 20–29 | 30–39 | 40–49 | 50–59 | 60–69 | 70–79 |
|---|---|---|---|---|---|---|---|---|
| *Frequency* | 8 | 32 | 142 | 216 | 240 | 206 | 143 | 13 |

**16.** The following data relate to the ages of a group of workers in a factory.

| *Ages* | *No. of workers* | *Ages* | *No. of workers* |
|---|---|---|---|
| 20—25 | 35 | 40—45 | 90 |
| 25—30 | 45 | 45—50 | 74 |
| 30—35 | 70 | 50—55 | 51 |
| 35—40 | 105 | 55—60 | 30 |

Draw the percentage cumulative curve and find from the graph the number of workers between the ages 28—48.

17. *(a)* The mean of marks obtained in an examination by a group of 100 students was found to be 49·96. The mean of the marks obtained in the same examination by another group of 200 students was 52·32. Find the mean of the marks obtained by both the groups of students taken together.

*(b)* A distribution consists of three components with frequencies 300, 200 and 600 having their means 16, 8 and 4 respectively. Find the mean of the combined distribution.

*(c)* The mean marks got by 300 students in the subject of Statistics are 45. The mean of the top 100 of them was found to be 70 and the mean of the last 100 was known to be 20. What is the mean of the remaining 100 students?

*(d)* The mean weight of 150 students in a certain class is 60 kilograms. The mean weight of boys in the class is 70 kilograms and that of the girls is 55 kilograms. Find the number of boys and number of girls in the class.

**Ans.** *(a)* 51·53, *(b)* 8, *(c)* 45, *(d)* Boys = 50, Girls = 100.

18. From the following data, calculate the percentage of workers getting wages

*(a)* more than Rs. 44, *(b)* between Rs. 22 and Rs. 58, *(c)* Find $Q_1$ and $Q_3$.

| *Wages (Rs.)* | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 | 50–60 | 60–70 | 70–80 |
|---|---|---|---|---|---|---|---|---|
| *No. of workers* | 20 | 45 | 85 | 160 | 70 | 55 | 35 | 30 |

**Hint.** Assuming that frequencies are uniformly distributed over the entire interval,

*(a)* Number of persons with wages more than Rs. 44 is

$$\left(\frac{50-44}{10}\times 70\right)+55+35+30=162$$

Hence the percentage of workers getting over Rs. 44 is

$$=\frac{162}{500}\times 100=32{\cdot}4\%$$

*(b)* Percentage of workers getting wages between Rs. 22 and Rs. 58 is

$$\left[\left(\frac{30-22}{10}\times 85\right)+160+70+\left(\frac{58-50}{10}\times 55\right)\right]\times 100 \div 500=68{\cdot}4\%$$

19. For the two frequency distributions give below the mean calculated from the first was 25·4 and that from the second was 32·5. Find the values of $x$ and $y$.

| *Class* | *Distribution I Frequency* | *Distribution II Frequency* |
|---|---|---|
| 10—20 | 20 | 4 |
| 20—30 | 15 | 8 |
| 30—40 | 10 | 4 |
| 40—50 | $x$ | $2x$ |
| 50—60 | $y$ | $y$ |

**Ans.** $x=3,\ y=2$

**20.** A number of particular articles has been classified according to their weights. After drying for two weeks the same articles have again been weighted and similarly classified. It is known that the median weight in the first weighing was 20·83 oz. while in the second weighing it was 17·35 oz. Some frequencies $a$ and $b$ in the first weighing and $x$ and $y$ in the second are missing. It is known that $a = \frac{1}{3}x$ and $b = \frac{1}{2}y$. Find out the values of the missing frequencies.

| *Class* | *Frequencies* 1st weighing | *Frequencies* IInd weighing | *Class* | *Frequencies* 1st weighing | *Frequencies* IInd weighing |
|---|---|---|---|---|---|
| 0—5 | $a$ | $x$ | 15—20 | 52 | 50 |
| 5—10 | $b$ | $y$ | 20—25 | 75 | 30 |
| 10—15 | 11 | 40 | 25—30 | 22 | 28 |

**Hint.** We have $x = 3a,\ y = 2b$,

$N_1 =$ Total frequency in 1st weighing $= 160 + a + b$.

$N_2 =$ Total frequency in 2nd weighing $= 148 + x + y = 148 + 3a + 2b$.

Using Median formula, we shall get

$$20{\cdot}83 = 20 + \frac{5}{75}\left[\frac{N_1}{2} - (63 + a + b)\right]$$

$$\Rightarrow \quad 15\,(20{\cdot}83 - 20) = \frac{160 + a + b}{2} - (63 + a + b)$$

$$\Rightarrow \quad 12{\cdot}45 = 17 - \frac{a + b}{2}$$

$$\Rightarrow \quad a + b = 2\,(17 - 12{\cdot}45) = 9{\cdot}10 \simeq 9 \qquad \ldots(*)$$

Since $a$ and $b$, being frequencies are integral valued, $a + b$ is also integral valued. Now the median of 2nd weighing gives:

$$17{\cdot}35 = 15 + \frac{5}{50}\left[\frac{148 + 3a + 2b}{2} - (40 + x + y)\right]$$

$$\Rightarrow \quad 10 \times 2{\cdot}35 = 74 + \frac{3a + 2b}{2} - 40 - 3a - 2b$$

$$\Rightarrow \quad \frac{3a + 2b}{2} = 34 - 23{\cdot}5 = 10{\cdot}5$$

$$\Rightarrow \quad 3a + 2b = 21 \qquad \ldots(**)$$

Multiplying $(*)$ by 3, we get

$$3a + 3b = 27 \qquad \ldots(***)$$

Subtracting $(**)$ from $(***)$, we get $b = 6$. Substituting in $(*)$, we get $a = 9 - 6 = 3$.

$$\therefore \quad a = 3,\ b = 6;\ \ x = 3a = 9,\ y = 2b = 12.$$

21. From the following table showing the wage distribution in a certain factory, determine :

(*a*) the mean wage,

(*b*) the median wage,

(*c*) the modal wage,

(*d*) the wage limits for the middle 50% of the wage earners,

(*e*) the percentages of workers who earned between Rs.75 and Rs.125.

(*f*) the percentage who earned more than Rs.150 per week, and

(*g*) the percentage who earned less than Rs.100 per week.

| *Weekly wages (Rs.)* | *No. of employees* | *Weekly wages (Rs.)* | *No. of employees* |
|---|---|---|---|
| 20–40 | 8 | 120–140 | 35 |
| 40–60 | 12 | 140–160 | 18 |
| 60–80 | 20 | 160–180 | 7 |
| 80–100 | 30 | 180–200 | 5 |
| 100–120 | 40 | | |

**Ans.** (*a*) $\overline{X} = 108{\cdot}5$, (*b*) Med. = 108·75, (*c*) Mo = 118·3, (*d*) 81·25 , 129·3 (*e*) 48, (*f*) 12, (*g*) 40.

22. (*a*) Explain how the ogives are drawn for any frequency distribution. Point out the method of finding out the values of median, mode, quartiles, deciles and percentiles graphically. Also, write down the formula for the computation of each of them for any frequency distribution.

(*b*) The following table gives the frequency distribution of marks in a class of 65 students.

| *Marks* | *No. of Students* | *Marks* | *No. of students* |
|---|---|---|---|
| 0–4 | 10 | 14–18 | 5 |
| 4–8 | 12 | 18–20 | 3 |
| 8–12 | 18 | 20–25 | 4 |
| 12–14 | 7 | 25 and over | 6 |
| *Total* | | | 65 |

Calculate : (*i*) Upper and lower quartiles.

(*ii*) No. of students who secured marks more than 17.

(*iii*) No. of students who secured marks between 10 and 15.

(*c*) The following table shows the age distribution of heads of families in a certain country during the year 1957. Find the median, the third quartile and the second decile of the distribution. Check your results by the graphical method.

| *Age of head of family years* | Under 25 | 25–29 | 30–34 | 35–44 | 45–54 | 55–64 | 65–74 | above 74 |
|---|---|---|---|---|---|---|---|---|
| *Number (million)* | 2·3 | 4·1 | 5·3 | 10·6 | 9·7 | 6·8 | 4·4 | 1·8 |
| | | | | | | | | Total 45 |

**Ans.** Md = 45·2 yrs.; $Q_3 = 57{\cdot}5$ yrs.; $D_2 = 32{\cdot}5$ yrs.

23. The following data represent travel expenses (other than transportation) for 7 trips made during November by a salesman for a small firm :

| *Trip* | *Days* | *Expense (Rs.)* | *Expense per day (Rs.)* |
|---|---|---|---|
| 1 | 0·5 | 13·50 | 27 |
| 2 | 2·0 | 12·00 | 6 |
| 3 | 3·5 | 17·50 | 5 |
| 4 | 1·0 | 9·00 | 9 |
| 5 | 9·0 | 27·00 | 3 |
| 6 | 0·5 | 9·00 | 18 |
| 7 | 8·5 | 17·00 | 2 |
| *Total* | 25·0 | 105·00 | 70 |

An auditor criticised these expenses as excessive, asserting that the average expense per day is Rs. 10 (Rs. 70 divided by 7). The salesman replied that the average is only Rs. 4·20 (Rs. 105 divided by 25) and that in any event the median is the appropriate measure and is only Rs. 3. The auditor rejoined that the arithmetic mean is the appropriate measure, but that the median is Rs. 6.

You are required to :

*(a)* Explain the proper interpretation of each of the four averages mentioned.

*(b)* Which average seems appropriate to you ?

24. *(a)* Define Geometric and Harmonic means and explain their uses in statistical analysis.

You take a trip which entails travelling 900 miles by train at an average speed of 60 m.p.h., 300 miles by boat at an average of 25 m.p.h., 400 miles by plane at 350 m.p.h. and finally 15 miles by taxi at 25 m.p.h. What is your speed for the entire distance?

*(b)* A train runs 25 miles at a speed of 30 m.p.h., another 50 miles at a speed of 40 m.p.h., then due to repairs of the track travels for 6 minutes at a speed of 10 m.p.h. and finally covers the remaining distance of 24 miles at a speed of 24 m.p.h. What is the average speed in m.p.h.?

*(c)* A man motors from $A$ to $B$. A large part of the distance is uphill and he gets a mileage of only 10 per gallon of gasoline. On the return trip he makes 15 miles per gallon. Find the harmonic mean of his mileage. Verify the fact that this is the proper average to be used by assuming that the distance from $A$ to $B$ is 60 miles.

*(d)* Calculate the average speed of a car running at the rate of 15 km.p.h. during the first 30 kms., at 20 km.p.h. during the second 30 kms. and at 25 kmp.h. during the third 30 kms.

25. The following table shows the distribution of 100 families according to their expenditure per week. Number of families corresponding to expenditure groups Rs. (10—20) and Rs.(30—40) are missing from the table. The median and

mode are given to be Rs.25 and 24 Calculate the missing frequencies and then arithmetic mean of the data :

| | | | | | |
|---|---|---|---|---|---|
| *Expenditure* : | 0—10 | 10—20 | 20—30 | 30—40 | 40—50 |
| *No. of families* : | 14 | ? | 27 | ? | 15 |

**Hint.**

| *Expenditure* | *No. of Families* | *Cumulative frequencies* |
|---|---|---|
| 0—10 | 14 | 14 |
| 10—20 | $f_1$ | $14+f_1$ |
| 20—30 | 27 | $41+f_1$ |
| 30—40 | $f_2$ | $41+f_1+f_2$ |
| 40—50 | 15 | $56+f_1+f_2$ |

$$\therefore \qquad 25=20+\frac{\dfrac{56+f_1+f_2}{2}-(14+f_1)}{27}\times 10$$

and
$$24=20+\frac{27-f_1}{2\times 27-f_1-f_2}\times 10$$

Simplying these equations, we get

$$f_1-f_2=1$$

and
$$3f_1-2f_2=27.$$

**Ans.** 25, 24

**26.** *(a)* The numbers 3·2, 5·8, 7·9 and 4·5 have frequencies $x$, $(x+2)$, $(x-3)$ and $(x+6)$ respectively. If their arithmetic mean is 4·876, find the value of $x$.

*(b)* If $M_{g.x}$ is the geometric mean of $N$ $x$'s amd $M_{g.y}$ is the geometric mean of $N$ $y$'s, then the geometric mean $M_g$ of the $2N$ values is given by

$$M_g^2=M_{g.x}\,M_{g.y}\,. \qquad \textbf{(Nagpur Univ. B.Sc., 1990)}$$

*(c)* The weighted geometric mean of the three numbers 229, 275 and 125 is 203. The weights for the first and the second numbers are 2 and 4 respectively. Find the weight of the third. **Ans.** 3.

**27.** The geometric mean of 10 observations on a certain variable was calculated as 16·2. It was later discovered that one of the observations was wrongly recorded as 12·9; in fact it was 21·9. Apply appropriate correction and calculate the correct geometric mean.

**Hint.** Correct value of the geometic mean, $G'$ is given by

$$G'=\left(\frac{(16\cdot 2)^{10}\times 21\cdot 9}{12\cdot 9}\right)^{1/10}=17\cdot 68$$

**28.** A variate takes the values $a, ar, ar^2, \ldots, ar^{n-1}$ each with frequency unity. If $A$, $G$ and $H$ are respectively the arithmetic mean, geometric mean and harmonic mean, show that

$$A = \frac{a(1-r^n)}{n(1-r)},\ G = ar^{(n-1)/2},\ H = \frac{an(1-r)r^{n-1}}{(1-r^n)}$$

Prove that $G^2 = AH$. Prove also that $A > G > H$ unless $r = 1$, when all the three means coincide.

**29.** If $\bar{x}_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\bar{x}_2 = \frac{1}{n}\sum_{i=2}^{n+1} x_i$ and $\bar{x}_3 = \frac{1}{n}\sum_{i=3}^{n+2} x_i$

then show that

(a) $\bar{x}_2 = \bar{x}_1 + \frac{1}{n}(x_{n+1} - x_1)$, and (b) $\bar{x}_3 = \bar{x}_2 + \frac{1}{n}(x_{n+2} - x_2)$

**30.** A distribution $x_1, x_2, \ldots, x_n$ with frequencies $f_1, f_2 \ldots, f_n$ transformed into the distribution $X_1, X_2, \ldots, X_n$ with the same corresponding frequencies by the relation $X_r = ax_r + b$, where $a$ and $b$ are constants. Show that the mean, median and mode of the new distribution are given in terms of those of the first distribution by the same transformation. **[Kanpur Univ. B.Sc., 1992]**

Use the method indicated above to find the mean of the following distribution: $x$ (duration of telephone conversation in seconds)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 49·5, | 149·5, | 249·5, | 349·5, | 449·5, | 549·5, | 649·5, | 749·5, | 849·5, | 949·5 |

$f$ (respective frequency)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 6 | 28 | 88 | 180 | 247 | 260 | 132 | 48 | 11 | 5 |

**31.** If $\bar{x}_w$ is the weighted mean of $x_i$'s with weights $w_i$, prove that

$$\left(\sum_{i=1}^{n} w_i\right)\left(\sum_{i=1}^{n} w_i(x_i - \bar{x}_w)^2\right) = \sum_{i=1}^{n}\sum_{j>i}^{n} w_i w_j (x_i - x_j)^2, \text{ where } \sum_{i=1}^{n} w_i \neq 0.$$

**(Allahabad Univ. B.Sc., 1992)**

**Hint.** $\left[\sum_{i=1}^{n}\sum_{j>i} w_i w_j (x_i - x_j)^2 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=i}^{n} w_i w_j (x_i - x_j)^2\right.$

$$\left. = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=i}^{n} w_i w_j \left\{\left(x_i - \bar{x}_w\right) - \left(x_j - \bar{x}_w\right)\right\}^2\right]$$

**32.** In a frequency table, the upper boundary of each class interval has a constant ratio to the lower boundary. Show that the geometric mean $G$ may be expressed by the formula :

$$\log G = x_0 + \frac{c}{N}\sum_i f_i\,(i-1)$$

where $x_0$ is the logarithm of the mid-value of the first interval and $c$ is the logarithm of the ratio between upper and lower boundaries.

**[Delhi Univ. B.Sc. (Stat. Hons.), 1990, 1986]**

33. Find the minimum value of :

(i) $f(x) = (x-6)^2 + (x+3)^2 + (x-8)^2 + (x+4)^2 + (x-3)^2$

(ii) $g(x) = |x-6| + |x+3| + |x-8| + |x+4| + |x-3|$.

[Delhi Univ. B.Sc.(Stat. Hons.), 1991]

**Hint.** The sum of squares of deviations is minimum when taken from arithmetic mean and the sum of absolute deviations is minimum when taken from median.

34. If $A$, $G$ and $H$ be the arithmetic mean, geometric mean and harmonic mean respectively of two positive numbers a and b, then prove that :

(i) $A \geq G \geq H$.

When does the equality sign hold?

(ii) $G^2 = AH$.

35 Calculate simple and weighted arithmetic averages from the following data and comment on them :

| *Designation* | *Monthly salary (in Rs.)* | *Strength of the cadre* |
|---|---|---|
| Class I Officers | 1,500 | 10 |
| Class II Officers | 800 | 20 |
| Subordinate staff | 500 | 70 |
| Clerical staff | 250 | 100 |
| Lower staff | 100 | 150 |

**Ans.** $\overline{X}$ = Rs. 630, $\overline{X}_w$ = Rs. 302·86. Latter is more representative.

36. Treating the number of letters in each word in the following passage as the variable $x$, prepare the frequency distribution table and obtain its mean, median, mode.

"The reliability of data must always be examined before any attempt is made to base conclusions upon them. This is true of all data, but particularly so of numerical data, which do not carry their quality written large on them. It is a waste of time to apply the refined theoretical methods of Statistics to data which are suspect from the beginning."

**Ans.** Mean = 4·565, Median = 4, Mode = 3.

## OBJECTIVE TYPE QUESTIONS

**I.** Match the correct parts to make a valid statement :

| | |
|---|---|
| (a) Arithmetic Mean | (i) $l + [f_2/(f_1 + f_2)] \times i$ |
| (b) Geometric Mean | (ii) $(x_1 . x_2 . \dots . x_n)^{1/n}$ |
| (c) Harmonic Mean | (iii) $\Sigma fX / \Sigma f$ |
| (d) Median | (iv) $l + \dfrac{N/2 - c.f.}{f} \times i$ |

(*e*) Mode (*v*) $\left[\frac{1}{n}\left(\frac{1}{x_1}+\frac{1}{x_2}+\ldots+\frac{1}{x_n}\right)\right]^{-1}$

(*vi*) $l=\frac{f_1-f_0}{2f_1-f_0-f_2}\times i$

**II.** Which measure of location will be suitable to compare:

(*i*) heights of students in two classes;

(*ii*) size of agricultural holdings;

(*iii*) average sales for various years;

(*iv*) intelligence of students;

(*v*) per capita income in several countries;

(*vi*) sale of shirts with collar size; 16″, $15\frac{1}{2}$″, 15″, 14″, 13″, 15″;

(*vii*) marks obtained 10, 8, 12, 4, 7, 11, and $X$ ($X < 5$).

**Ans.** (*i*) Mean, (*ii*) Mode, (*iii*) Mean, (*iv*) Median, (*v*) Mean, (*vi*) Mode, (*vii*) Median.

**III.** Which of the following are true for all sets of data?

(*i*) Arithmetic Mean ≤ median ≤ mode,

(*ii*) Arithmetic mean ≥ median ≥ mode,

(*iii*) Arithmetic mean = median = mode

(*iv*) None of these.

**IV.** Which of the following are true in respect of any distribution?

(*i*) The percentile points are in the ascending order.

(*ii*) The percentile points are equispaced.

(*iii*) The median is the mid-point of the range and the distribution.

(*iv*) A unique median value exists for each and every distribution.

**V.** Find out the missing figures :

(*a*) Mean = ? (3 Median – Mode).

(*b*) Mean – Mode = ? (Mean – Median).

(*c*) Median = Mode + ? (Mean – Mode).

(*d*) Mode = Mean – ? (Mean – Median).

**Ans.** (*a*) 1/2, (*b*) 3, (*c*) 2/3, (*d*) 3.

**VI.** Fill in blanks :

(*i*) Harmonic mean of a number of observations is ......

(*ii*) The geometric mean of 2, 4, 16, and 32 is ......

(*iii*) The strength of 7 colleges in a city are 385; 1,748; 1,343; 1,935; 786; 2,874 and 2,108. Then the median strength is ......

(*iv*) The geometric mean of a set of values lies between arithmetic mean and...

(*v*) The mean and median of 100 items are 50 and 52 respectively. The value of the largest item is 100. It was later found that it is actually 110. Therefore, the true mean is ... and the true median is ......

*(vi)* The algebraic sum of the deviations of 20 observations measured from 30 is 2. Therefore, mean of these observations is .....

*(vii)* The relationship between A.M., G.M. and H.M. is .....

*(viii)* The mean of 20 observations is 15. On checking it was found that two observations were wrongly copied as 3 and 6. If wrong observations are replaced by correct values 8 and 4, then the correct mean is .....

*(ix)* Median = ..... Quartile.

*(x)* Median is the average suited for ..... classes.

*(xi)* A distribution with two modes is called ..... and with more than two modes is called .....

*(xii)* ..... is not affected by extreme observations.

**Ans.** *(ii)* 8 ; *(iii)* 1,748 ; *(iv)* H.M. ; *(v)* 50·1, 52 ; *(vi)* 30·1 ; *(vii)* A.M. ≥ G.M. ≥ H.M. ; *(viii)* 15·15 ; *(ix)* Second ; *(x)* Open end ; *(xi)* Bimodal, multimodal ; *(xii)* Median or mode.

**VII.** For the questions given below, give correct answers.

*(i)* The algebraic sum of the deviations of a set of $n$ values from their arithmetic mean is

*(a)* $n$, *(b)* 0, *(c)* 1, *(d)* none of these.

*(ii)* The most stable measure of central tendency is

*(a)* the mean, *(b)* the median, *(c)* the mode, *(d)* none of these.

*(iii)* 10 is the mean of a set of 7 observations and 5 is the mean of a set of 3 observations. The mean of a combined set is given by

*(a)* 15, *(b)* 10, *(c)* 8·5, *(d)* 7·5, *(e)* none of these.

*(iv)* The mean of the distribution, in which the value of $x$ are 1, 2, ..., $n$, the frequency of each being unity is:

*(a)* $n(n+1)/2$, *(b)* $n/2$, *(c)* $(n+1)/2$, *(d)* none of these.

*(v)* The arithmetic mean of the numbers 1, 2, 3, ..., $n$ is

*(a)* $\dfrac{n(n+1)(2n+1)}{6}$, *(b)* $\dfrac{n(n+1)^2}{4}$, *(c)* $\dfrac{n(n+1)}{2}$, *(d)* none of these.

*(ii)* The most stable measure of central tendency

*(vi)* The point of intersection of the 'less than' and the 'greater than' ogive corresponds to

*(a)* the mean, *(b)* the median, *(c)* the geometric mean, *(d)* none of these.

*(vii)* When $x_i$ and $y_i$ are two variables ($i = 1, 2, ..., n$) with G.M.'s $G_1$ and $G_2$ respectively then the geometric mean of $\left(\dfrac{x_i}{y_i}\right)$ is

*(a)* $\dfrac{G_1}{G_2}$, *(b)* antilog $\left(\dfrac{G_1}{G_2}\right)$, *(c)* $n(\log G_1 - \log G_2)$;

(d) Antilog $\left( \frac{\log G_1 - \log G_2}{2n} \right)$

**Ans.** *(i)* (b) ; *(ii)* (a) ; *(iii)* (c) ; *(iv)* (c) ; *(v)* (d) ; *(vi)* (b) ; *(vii)* (a).

**VIII. State which of the following statements are True and which are False. In case of false statements give the correct statement.**

*(i)* The harmonic mean of $n$ numbers is the reciprocal of the Arithmetic mean of the reciprocals of the numbers.

*(ii)* For the wholesale manufacturers interested in the type which is usually in demand, median is the most suitable average.

*(iii)* The algebraic sum of the deviations of a series of individual observations from their mean is always zero.

*(iv)* Geometric mean is the appropriate average when emphasis in on the rate of change rather than the amount of change.

*(v)* Harmonic mean becomes zero when one of the items is zero.

*(vi)* Mean lies between median and mode.

*(vii)* Cumulative frequency is not-decreasing.

*(viii)* Geometric mean is the arithmetic mean of harmonic mean and arithmetic mean.

*(ix)* Mean, median mode have the same unit.

*(x)* One quintal of wheat was purchased at 0·8 kg. per rupee and another quintal at 1·2 kg. per rupee. The average rate per rupee is 1kg.

*(xi)* One limitation of the median is that it cannot be calculated from a frequency distribution with open end classes.

*(xii)* The arithmetic mean of a frequency distribution is always located in the class which has the greatest number of frequencies.

*(xiii)* In a moderately asymmetrical distribution, the mean, median and mode are the same.

*(xiv)* It is really immaterial in which class an item falling at the boundary between two classes is listed.

*(xv)* The median is not affected by extreme items.

*(xvi)* The median is the point about which the sum of squared deviations is minimum.

(xvii) In construction of the frequency distribution, the selection of the class interval is arbitrary.

*(xviii)* Usual attendance of B.Sc. class is 35 per day. So for 100 working days total attendance is 3,500.

*(xix)* A car travels 100 miles at a speed of 40 m.p.h. and another 400 miles at a speed of 30 m.p.h. So the average speed for the whole journey is either 35 m.p.h. or 33 m.p.h.

*(xx)* In calculating the mean for grouped data, the assumption is made that the mean of the items in each class is equal to the mid-value of the class.

*(xxi)* The geometric mean of a group of numbers is less than the arithmetic mean in all cases, except in the special case in which the numbers are all the same.

*(xxii)* The geometric mean equals the antilog of the arithmetic mean of the logs of the values.

*(xxiii)* The median may be considered more typical than the mean because the median is not affected by the size of the extremes.

*(xxiv)* The Harmonic Mean of a series of fractions is the same as the reciprocal of the arithmetic mean of the series.

*(xxv)* In a frequency distribution the true value of mode cannot be calculated exactly.

IX. In each of the following cases, explain whether the description applies to mean, median or both.

*(i)* it can be calculated from a frequency distribution with open-end classes.

*(ii)* the values of all items are taken into consideration in the calculation.

*(iii)* the values of extreme items do not influence the average.

*(iv)* In a distribution with a single peak and moderate skewness to the right it is closer to the concentration of the distribution.

Ans. *(i)* median, *(ii)* mean, *(iii)* median, *(iv)* median.

X. Be brief in your answer :

*(a)* The production in an industrial unit was 10,000 units during 1981 and in 1980 the production was 25,000 units. Hence the production has declined by 150 percent. Comment.

*(b)* A man travels by a car for 4 days. He travelled for 10 hours each day. He drove on the first day at the rate of 45 km per hour, second day at 40 km. per hour, third day at the rate of 38 km. per hour and the fourth day at the rate of 37 km. per hour.

Which average, harmonic mean or arithmetic mean or median will give us his average speed? Why?

*(c)* It is seen from records that a country does not export more than 5 % of its total production. Hence export trade is not vital to the economy of that country. Is the conclusion right?

*(d)* A survey revealed that the children of engineers, doctors and lawyers have high intelligence quotients. It further revealed that the grandfathers of these children were also highly intelligent. Hence the inference is that intelligence is hereditary. Do you agree?

XI. Do you agree with the following interpretations made on the basis of the facts given. Explain briefly your answer.

*(a)* The number of deaths in military in the recent war was 10 out of 1,000 while the number of deaths in Hyderabad in the same period was 18 per 1,000. Hence it is safe to join military service than to live in the city of Hyderabad.

*(b)* The examination result in a college $X$ was 70% in the year 1991. In the same year and at the same examination only 500 out of 750 students were successful in college $Y$. Hence the teaching standard in college $X$ was better.

*(c)* The average daily production in a small-scale factory in January 1991 was 4,000 candles and 3,800 candles in February 1981. So the workers were more efficient in January.

*(d)* The increase in the price of a commodity was 25%. Then the price decreased by 20% and again increased by 10%. So the resultant increase in the price was $25 - 20 + 10 = 15\%$.

*(e)* The rate of tomato in the first week of January was 2 kg. for a rupee and in the 2nd week was 4 kg. for a rupee. Hence the average price of tomato is $\frac{1}{2}(2 + 4) = 3$ kg. for a rupee.

XII. *(a)* The mean mark of 100 students was given to be 40. It was found later that a mark 53 was read as 83. What is the corrected mean mark?

*(b)* The mean salary paid to 1,000 employees of an establishment was found to be Rs. 108·40. Later on, after disbursement of salary it was discovered that the salary of two employees was wrongly entered as Rs. 297 and Rs. 165. Their correct salaries were Rs. 197 and Rs. 185. Find the correct arithmetic mean.

*(c)* Twelve persons gambled on a certain night. Seven of them lost at an average rate of Rs. 10·50 while the remaining five gained at an average of Rs. 13·00. Is the information given above correct? If not, why?

CHAPTER THREE

# *Measures of Dispersion, Skewness and Kurtosis*

**3·1. Dispersion.** Averages or the measures of central tendency give us an idea of the concentration of the observations about the central part of the distribution. If we know the average alone we cannot form a complete idea about the distribution as will be cear from the following example.

Consider the series *(i)* 7, 8, 10, 11, *(ii)* 3, 6, 9, 12, 15, *(iii)* 1, 5, 9, 13, 17. In all these cases we see that n, the number of observations is 5 and the mean is 9. If we are given that the mean of 5 observations is 9, we cannot form an idea as to whether it is the average of first series or second series or third series or of any other series of 5 observations whose sum is 45. Thus we see that the measures of central tendency are inadequate to give us a complete idea of the distribution. They must be supported and supplemented by some other measures. One such measure is *Dispersion.*

Literal meaning of dispersion is *'scatteredness'*. We study dispersion to have an idea about the homogeneity or heterogeneity of the distribution. In the above case we say that series *(i)* is more homogeneous (less dispersed) than the series *(ii)* or *(iii)* or we say that series *(iii)* is more heterogeneous (more scattered) than the series *(i)* or *(ii)*.

**3·2. Characteristics for an Ideal Measure of Dispersion.** The desiderata for an ideal measure of dispersion are the same as those for an ideal measure of central tendency, *viz.,*

*(i)* It should he rigidly defined.

*(ii)* It should be easy to calculate and easy to understand.

*(iii)* It should be based on all the observations.

*(iv)* It should be amenable to further mathematical treatment.

*(v)* It should be affected as little as possible by fluctuations of sampling.

**3·3. Measures of Dispersion.** The following are the measures of dispersion:

*(i) Range,*

*(ii) Quartile deviation or Semi-interquartile range,*

*(iii) Mean deviation, and*

*(iv) Standard deviation.*

**3·4. Range.** The range is the difference between two extreme obsertations of the distribution. If $A$ and $B$ are the greatest and smallest observations respectively in a distribution, then its range is $A - B$.

Range is the simplest but a crude measure of dispersion. Since it is based on two extreme observations which themselves are subject to chance fluctuations, it is not at all a reliable measure of dispersion.

**3·5. Quartile Deviation.** Quartile deviation or semi-interquartile range

$Q$ is given by

$$Q = \frac{1}{2}(Q_3 - Q_1), \qquad ...(3\cdot 1)$$

where $Q_1$ and $Q_3$ are the first and third quartiles of the distribution respectively.

Quartile deviation is definitely a better measure than the range as it makes use of 50% of the data. But since it ignores the other 50% of the data, it cannot be regarded as a reliable measure.

**3·6. Mean Deviation.** If $x_i \mid f_i$, $i = 1, 2, ..., n$ is the frequency distribution, then mean deviation from the average $A$, (usually mean, median or mode), is given by

$$\text{Mean deviation} = \frac{1}{N} \sum_i f_i \mid x_i - A \mid, \ \Sigma f_i = N \qquad ...(3\cdot 2)$$

where $\mid x_i - A \mid$ represents the modulus or the absolute value of the deviation $(x_i - A)$, when the –ive sign is ignored.

Since mean deviation is based on all the observations, it is a better measure of dispersion than range or quartile deviation. But the step of ignoring the signs of the deviations $(x_i - A)$ creates artificiality and renders it useless for further mathematical treatment.

It may be pointed out here that mean deviation is least when taken from median. (The proof is given for continuous variable in Chapter 5)

**3·7. Standard Deviation and Root Mean Square Deviation.** Standard deviation, usually denoted by the Greek letter small sigma ($\sigma$), is the positive square root of the arithmetic mean of the squares of the deviations of the given values from their arithmetic mean. For the frequency distribution $x_i \mid f_i$, $i = 1, 2, ..., n$,

$$\sigma = \sqrt{\frac{1}{N} \sum_i f_i (x_i - \bar{x})^2} \qquad ...(3\cdot 3)$$

where $\bar{x}$ is the arithmetic mean of the distribution and $\sum_i f_i = N$.

The step of squaring the deviations $(x_i - \bar{x})$ overcomes the drawback of ignoring the signs in mean deviation. Standard deviation is also suitable for further mathematical treatment (§ 3·7·3). Moreover of all the measures, standard deviation is affected least by fluctuations of sampling.

Thus we see that standard deviation satisfies almost all the properties laid down for an ideal measure of dispersion except for the general nature of extracting the square root which is not readily comprehensible for a non-mathematical person. It may also be pointed out that standard deviation gives greater weight to extreme values and as such has not found favour with economists or businessmen who are more interested in the results of the modal class. Taking into consideration the pros and cons and also the wide applications of standard deviation in statistical theory, we may regard standard deviation as the best and the most powerful measure of dispersion.

The square of standard deviation is called the *variance* and is given by

$$\sigma^2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2 \qquad \text{...(3·3a)}$$

Root mean square deviation, denoted by '$s$' is given by

$$s = \sqrt{\frac{1}{N} \sum_i f_i (x_i - A)^2} \qquad \text{...(3·4)}$$

where $A$ is any arbitrary number. $s^2$ is called mean square deviation.

**3·7·1. Relation between $\sigma$ and $s$.** By definition, we have

$$s^2 = \frac{1}{N} \sum_i f_i (x_i - A)^2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x} + \bar{x} - A)^2$$

$$= \frac{1}{N} \sum_i f_i \left[ (x_i - \bar{x})^2 + (\bar{x} - A)^2 + 2 (\bar{x} - A)(x_i - \bar{x}) \right]$$

$$= \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2 + (\bar{x} - A)^2 \frac{1}{N} \sum_i f_i + 2 (\bar{x} - A) \sum_i f_i (x_i - \bar{x}),$$

$(\bar{x} - A)$, being constant is taken outside the summation sign. But $\sum_i f_i (x_i - \bar{x}) = 0$, being the algebraic sum of the deviations of the given values from their mean. Thus

$$s^2 = \sigma^2 + (\bar{x} - A)^2 = \sigma^2 + d^2, \text{ where } d = \bar{x} - A$$

Obviously $s^2$ will be least when $d = 0$, *i.e.*, $\bar{x} = A$. Hence mean square deviation and consequently root mean square deviation is least when the deviations are taken from $A = \bar{x}$, *i.e.*, standard deviation is the least value of root mean square deviation.

The same result could be obtained alternatively as follows:
Mean square deviation is given by

$$s^2 = \frac{1}{N} \sum_i f_i (x_i - A)^2$$

It has been shown in § 2·5·1 Property 2 that $\sum_i f_i (x_i - A)^2$ is minimum when $A = \bar{x}$. Thus mean square deviation is minimum when $A = \bar{x}$ and its minimum value is

$$(s^2)\min = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2 = \sigma^2$$

*Hence variance is the minimum value of mean square deviation or standard deviation is the minimum value of root mean square deviation.*

**3·7·2. Different Formulae For Calculating Variance.** By definition, we have

$$\sigma^2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2$$

More precisely we write it as $\sigma_x^2$, *i.e.*, variance of $x$. Thus

$$\sigma_x^2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2 \qquad ...(3\cdot5)$$

If $\bar{x}$ is not a whole number but comes out to be in fractions, the calculation of $\sigma_x^2$ by using (3·5) is very cumbersome and time consuming. In order to overcome this difficulty, we shall develop different forms of the formula (3·5) which reduce the arithmetic to a great extent and are very useful for computational work. In the following sequence the summation is extended over $i$ from 1 to $n$.

$$\sigma_x^2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2 = \frac{1}{N} \sum_i f_i (x_i^2 + \bar{x}^2 - 2\, x_i \bar{x})$$

$$= \frac{1}{N} \sum_i f_i x_i^2 + \bar{x}^2 \frac{1}{N} \sum_i f_i - 2\bar{x} \cdot \frac{1}{N} \sum_i f_i x_i$$

$$= \frac{1}{N} \sum_i f_i x_i^2 + \bar{x}^2 - 2\bar{x}^2 = \frac{1}{N} \sum_i f_i x_i^2 - \bar{x}^2 \qquad ...(3\cdot6)$$

$$\Rightarrow \qquad \sigma_x^2 = \frac{1}{N} \sum_i f_i x_i^2 - \left( \frac{1}{N} \sum_i f_i x_i \right)^2 \qquad ...(3\cdot6a)$$

If the values of $x$ and $f$ are large the calculation of $fx$, $fx^2$ is quite tedious. In that case we take the deviations from any arbitrary point '$A$'. Generally the point in the middle of the distribution is much convenient though the formula is true in general. We have

$$\sigma_x^2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2 = \frac{1}{N} \sum_i f_i (x_i - A + A - \bar{x})^2$$

$$= \frac{1}{N} \sum_i f_i (d_i + A - \bar{x})^2, \text{ where } d_i = x_i - A.$$

$$\sigma_x^2 = \frac{1}{N} \sum_i f_i [\, d_i^2 + (A - \bar{x})^2 + 2\,(A - \bar{x})\, d_i\,]$$

$$= \frac{1}{N} \sum_i f_i d_i^2 + (A - \bar{x})^2 + 2\,(A - \bar{x}) \cdot \frac{1}{N} \sum_i f_i d_i$$

We know that if $d_i = x_i - A$ then $\bar{x} = A + \frac{1}{N} \sum_i f_i d_i$

$$\therefore \qquad A - \bar{x} = -\frac{1}{N} \sum_i f_i d_i$$

Hence

$$\sigma_x^2 = \frac{1}{N} \sum f_i d_i^2 + \left( -\frac{1}{N} \sum_i f_i d_i \right)^2 + 2 \left( -\frac{1}{N} \sum_i f_i d_i \right) \left( \frac{1}{N} \sum_i f_i d_i \right)$$

$$= \frac{1}{N} \sum_i f_i d_i^2 - \left( \frac{1}{N} \sum_i f_i d_i \right)^2 \qquad ...(3\cdot7)$$

$$\Rightarrow \qquad \sigma_x^2 = \sigma_d^2 \qquad \text{[On comparison with (3·6a)]}$$

Hence variance and consequently standard deviation is independent of change of origin.

If we take $d_i = (x_i - A)/h$ so that $(x_i - A) = hd_i$, then

$$\sigma_x^2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2 = \frac{1}{N} \sum_i f_i (x_i - A + A - \bar{x})^2$$

$$= \frac{1}{N} \sum_i f_i (hd_i + A - \bar{x})^2$$

$$= h^2 \frac{1}{N} \sum_i f_i d_i^2 + (A - \bar{x})^2 + 2(A - \bar{x}) . h . \frac{1}{N} \sum_i f_i d_i$$

Using $\bar{x} = A + h \dfrac{\Sigma f_i d_i}{N}$, we get

$$\sigma_x^2 = h^2 \left[ \frac{1}{N} \sum_i f_i d_i^2 - \left( \frac{1}{N} \sum_i f_i d_i \right)^2 \right] = h^2 \sigma_d^2, \qquad \text{...(3·8)}$$

which shows that *variance is not independent of change of scale.*

**Aliter.** If $d_i = \dfrac{x_i - A}{h}$, then

$$x_i = A + hd_i \quad \text{and} \quad \bar{x} = A + h . \frac{1}{N} \sum_i f_i d_i = A + h \bar{d}$$

Obviously $\qquad x_i - \bar{x} = h(d_i - \bar{d})$

$$\therefore \qquad \sigma_x^2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2 = h^2 . \frac{1}{N} \sum_i f_i (d_i - \bar{d})^2 = h^2 \sigma_d^2$$

Hence variance is independent of change of origin but not of scale.

**Example 3·1.** *Calculate the mean and standard deviation for the following table giving the age distribution of 542 members.*

| *Age in years :* | 20—30 | 30—40 | 40—50 | 50—60 | 60—70 | 70—80 | 80—90 |
|---|---|---|---|---|---|---|---|
| *No. of members :* | 3 | 61 | 132 | 153 | 140 | 51 | 2 |

**Solution.** Here we take $d = \dfrac{x - A}{h} = \dfrac{x - 55}{10}$

| *Age group* | *Mid-value* $(x)$ | *Frequncy* $(f)$ | $d = \dfrac{x - 55}{10}$ | $fd$ | $fd^2$ |
|---|---|---|---|---|---|
| 20 — 30 | 25 | 3 | −3 | −9 | 27 |
| 30 — 40 | 35 | 61 | −2 | −122 | 244 |
| 40 — 50 | 45 | 132 | −1 | −132 | 132 |
| 50 — 60 | 55 | 153 | 0 | 0 | 0 |
| 60 — 70 | 65 | 140 | 1 | 140 | 140 |
| 70 — 80 | 75 | 51 | 2 | 102 | 204 |
| 80 — 90 | 85 | 2 | 3 | 6 | 18 |
| | | $N = \Sigma f = 542$ | | $\Sigma fd = -15$ | $\Sigma fd^2 = 765$ |

$$\bar{x} = A + h \frac{\Sigma fd}{N} = 55 + \frac{10 \times (-15)}{542} = 55 - 0·28 = 54·72 \text{ years.}$$

$$\sigma^2 = h^2 \left[ \frac{1}{N} \Sigma fd^2 - \left( \frac{1}{N} \Sigma fd \right)^2 \right] = 100 \left[ \frac{765}{542} - (0·28)^2 \right]$$

$= 100 \times 1{\cdot}333 = 133{\cdot}3$

$\therefore$ $\sigma$ (standard deviation) = 11·55 years

**Example 3·2.** *Prove that for any discrete distribution standard deviation is not less than mean deviation from mean.*

**[Delhi Univ. B.Sc. (Stat. Hons.), 1989]**

**Solution.** Let $x_i \mid f_i$, $i = 1, 2, 3, ..., n$ be any discrete distribution. Then we have to prove that

$$\text{S.D.} \nless \text{Mean deviation from mean}$$

$$\Rightarrow \quad (\text{S.D.})^2 \nless (\text{Mean deviation from mean})^2$$

$$\Rightarrow \quad (\text{S.D.})^2 \geq (\text{M. D. from mean})^2$$

$$\Rightarrow \quad \frac{1}{N} \sum_{i=1}^{n} f_i (x_i - \bar{x})^2 \geq \left( \frac{1}{N} \sum_{i=1}^{n} f_i \mid x_i - \bar{x} \mid \right)^2$$

If we put $\mid x_i - \bar{x} \mid = z_i$, then we have to prove that

$$\frac{1}{N} \sum_{i=1}^{n} f_i z_i^2 \geq \left( \frac{1}{N} \sum_{i=1}^{n} f_i z_i \right)^2$$

*i.e.,* $$\frac{1}{N} \sum_{i=1}^{n} f_i z_i^2 - \left( \frac{1}{N} \sum_{i=1}^{n} f_i z_i \right)^2 \geq 0$$

*i.e.,* $$\frac{1}{N} \sum_{i=1}^{n} f_i (z_i - \bar{z})^2 \geq 0$$

*i.e.,* $$\sigma_z^2 \geq 0,$$

which is always true. Hence the result.

**Example 3·3.** *Find the mean deviation from the mean and standard deviation of A.P. $a, a + d, a + 2d, ..., a + 2nd$ and verify that the latter is greater than the former.* **[Delhi Univ. B.Sc. (Stat. Hons.), 1990]**

**Solution.** We know that the mean of a series in A.P. is the mean of its first and last term. Hence the mean of the given series is

$$\bar{x} = \frac{1}{2}(a + a + 2nd) = a + nd$$

| $x$ | $\mid x - \bar{x} \mid$ | $(x - \bar{x})^2$ |
|---|---|---|
| $a$ | $nd$ | $n^2d^2$ |
| $a + d$ | $(n-1)d$ | $(n-1)^2 d^2$ |
| $a + 2d$ | $(n-2)d$ | $(n-2)^2 d^2$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $a + (n-2)d$ | $2d$ | $2^2 . d^2$ |
| $a + (n-1)d$ | $d$ | $1^2 . d^2$ |
| $a + nd$ | $0$ | $0$ |
| $a + (n+1)d$ | $d$ | $1^2 . d^2$ |
| $a + (n+2)d$ | $2d$ | $2^2 . d^2$ |

| | | |
|---|---|---|
| $\vdots$ | $\vdots$ | $\vdots$ |
| $a + (2n - 2)\,d$ | $(n - 2)\,d$ | $(n - 2)^2 d^2$ |
| $a + (2n - 2)\,d$ | $(n - 1)\,d$ | $(n - 2)^2 d^2$ |
| $a + 2nd$ | $nd$ | $n^2 d^2$ |

$$\text{Mean deviation from mean} = \frac{1}{2n + 1}\,\Sigma\,|x - \bar{x}|$$

$$= \frac{1}{2n + 1}\,2.d\,(1 + 2 + 3 + \ldots + n)$$

$$= \frac{n\,(n + 1)d}{(2n + 1)}$$

$$\sigma^2 = \frac{1}{2n + 1}\,\Sigma\,(x - \bar{x})^2 = \frac{1}{2n + 1}\,2.d^2\,(1^2 + 2^2 + 3^2 + \ldots + n^2)$$

$$= \frac{1}{2n + 1}\,2d^2 \cdot \frac{n\,(n + 1)\,(2n + 1)}{6} = \frac{n\,(n + 1)d^2}{3}$$

Hence standard deviation

$$\sigma = \sqrt{\frac{n\,(n + 1)}{3}} \times d$$

**Verification.**

$$\text{S.D.} > \text{M.D. from mean}$$

if $\quad (\text{S.D.})^2 > (\text{M.D. from mean})^2$

*i.e.,* if $\quad \dfrac{n\,(n + 1)d^2}{3} > \left(\dfrac{n\,(n + 1)d}{2n + 1}\right)^2$

or if $\quad (2n + 1)^2 > 3n\,(n + 1)$

or if $\quad n^2 + n + 1 > 0$

or if $\quad (n + \frac{1}{2})^2 + \frac{3}{4} > 0$

which is always ture.

**Example 3·4.** *Show that in a discrete series if deviations are small compared with mean M so that $(x/M)^3$ and higher powers of $(x/M)$ are neglected, we have*

$$(i)\ G = M\left(1 - \frac{1}{2} \cdot \frac{\sigma^2}{M^2}\right),$$

$$(ii)\ M^2 - G^2 = \sigma^2, \text{ and } (iii)\ H = M\left(1 - \frac{\sigma^2}{M^2}\right),$$

*where M is the arithemetic mean, G, the grometric mean, H, the harmonic mean and σ is the standard deviation of the distribution.*

**Solution.** Let $X_i\,|\,f_i$, $i = 1, 2, \ldots, n$ be the given frequency distribution. Then we are given that $x_i = X_i - M$, *i.e.,* $X_i = x_i + M$ where $M$ is the mean of the distribution. We have

$$\sum_i f_i x_i = \sum_i f_i (X_i - M) = 0, \qquad ...(1)$$

being the algebraic sum of the deviations of the given values from their mean. Also

$$\sum_i f_i x_i^2 = \sum_i f_i (X_i - M)^2 = \sigma^2 \qquad ...(2)$$

*(i)* By definition, we have

$$G = (X_1^{f_1} . X_2^{f_2} \ldots X_n^{f_n})^{1/N}, \text{ where } N = \Sigma f_i$$

$$\log G = \frac{1}{N} \sum_i f_i \log X_i = \frac{1}{N} \sum_i f_i \log (x_i + M)$$

$$= \frac{1}{N} \sum_i \left\{ f_i \log \left[ M \left( 1 + \frac{x_i}{M} \right) \right] \right\}$$

$$= \frac{1}{N} \sum_i \left\{ f_i \left[ \log M + \log \left( 1 + \frac{x_i}{M} \right) \right] \right\}$$

$$= \log M + \frac{1}{N} \sum_i f_i \log \left( 1 + \frac{x_i}{M} \right)$$

$$= \log M + \frac{1}{N} \sum_i f_i \left[ \frac{x_i}{M} - \frac{1}{2} \frac{x_i^2}{M^2} + \frac{1}{3} \left( \frac{x_i}{M} \right)^3 + \ldots \right],$$

the expansion of $\log \left( 1 + \dfrac{x_i}{M} \right)$ in ascending powers of $(x_i/M)$ being valid since $| x_i / M | < 1$. Neglecting $(x_i/M)^3$ and higher powers of $(x_i/M)$, we get

$$\log G = \log M + \frac{1}{NM} \sum_i f_i x_i - \frac{1}{2M^2} \cdot \frac{1}{N} \sum_i f_i x_i^2$$

$$= \log M - \frac{\sigma^2}{2M^2}, \qquad \text{[On using (1) and (2)]}$$

$$= \log \left( M e^{-\sigma^2/2M^2} \right)$$

$$\Rightarrow \qquad G = M e^{-\sigma^2/2M^2} = M \left( 1 - \frac{\sigma^2}{2M^2} \right),$$

neglecting higher powers.

Hence $\quad G = M \left( 1 - \dfrac{1}{2} \cdot \dfrac{\sigma^2}{M^2} \right)$. $\qquad ...(3)$

*(ii)* Squaring both sides in (3), we get

$$G^2 = M^2 \left( 1 - \frac{1}{2} \cdot \frac{\sigma^2}{M^2} \right)^2 = M^2 \left( 1 - \frac{\sigma^2}{M^2} \right) = M^2 - \sigma^2,$$

neglecting $(\sigma/M)^4$.

$$\therefore \qquad M^2 - G^2 = \sigma^2 \qquad ...(4)$$

*(iii)* By definition, harmonic mean $H$ is given by

$$\frac{1}{H}=\frac{1}{N}\sum_i (f_i/X_i)=\frac{1}{N}\sum_i [f_i/(x_i+M)]$$

$$=\frac{1}{MN}\sum_i \frac{f_i}{[1+(x_i/M)]}=\frac{1}{MN}\sum_i f_i\left(1+\frac{x_i}{M}\right)^{-1}$$

Since $\left|\dfrac{x_i}{M}\right|<1$, the expansion of $\left(1+\dfrac{x_i}{M}\right)^{-1}$ in ascending powers of $(x_i/M)$ is valid. Neglecting $(x_i/M)^3$ and higher powers of $(x_i/M)$, we get

$$\frac{1}{H}=\frac{1}{MN}\sum_i f_i\left(1-\frac{x_i}{M}+\frac{x_i^2}{M^2}\right)$$

$$=\frac{1}{M}\left(\frac{1}{N}\sum_i f_i-\frac{1}{MN}\sum_i f_i x_i+\frac{1}{M^2}\frac{1}{N}\sum_i f_i x_i^2\right)$$

$$=\frac{1}{M}\left(1+\frac{\sigma^2}{M^2}\right) \qquad \text{[On using (1) and (2)]}$$

$$\therefore \qquad H=M\left(1+\frac{\sigma^2}{M^2}\right)^{-1}=M\left(1-\frac{\sigma^2}{M^2}\right),$$

higher powers being neglected.

Hence $$H=M\left(1-\frac{\sigma^2}{M^2}\right) \qquad \text{...(5)}$$

**Example 3·5.** *For a group of 200 candidates, the mean and standard deviation of scores were found to be 40 and 15 respectively. Later on it was discovered that the scores 43 and 35 were misread as 34 and 53 respectively. Find the corrected mean and standard deviation corresponding to the corrected figures.*

**Solution.** Let $x$ be the given variable. We are given $n = 200$, $\bar{x} = 40$ and $\sigma = 15$

Now $$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i \quad\Rightarrow\quad \sum_i x_i=n\bar{x}=200\times 40=8000$$

Also $$\sigma^2=\frac{1}{n}\sum_i x_i^2-\bar{x}^2$$

$$\therefore \qquad \sum_i x_i^2=n(\sigma^2+\bar{x}^2)=200\,(225+1600)=365000$$

$$\text{Corrected } \sum_i x_i=8000-34-53+43+35=7991$$

and $$\text{Corrected } \sum_i x_i^2=365000-(34)^2-(53)^2+(43)^2+(35)^2=364109$$

Hence, $$\text{Corrected mean}=\frac{7991}{200}=39{\cdot}955$$

$$\text{Corrected } \sigma^2 = \frac{364109}{200} - (39{\cdot}955)^2 = 1820{\cdot}54 - 1596{\cdot}40 = 224{\cdot}14$$

$\therefore$ Corrected standard deviation = 14·97

**3·7·3. Theorem.** *(Variance of the combined series). If $n_1, n_2$ are the sizes; $\bar{x}_1, \bar{x}_2$ the means, and $\sigma_1, \sigma_2$ the standard deviations of two series, then the standard deviation $\sigma$ of the combined series of size $n_1 + n_2$ is given by*

$$\sigma^2 = \frac{1}{n_1 + n_2}\left[n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)\right] \qquad \ldots(3{\cdot}9)$$

*where* $d_1 = \bar{x}_1 - \bar{x}$, $d_2 = \bar{x}_2 - \bar{x}$

*and* $\bar{x} = \dfrac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$, *is the mean of the combined series.*

**Proof.** Let $x_{1i}$; $i = 1, 2, \ldots, n_1$ and $x_{2j}$; $j = 1, 2, \ldots, n_2$, be the two series then

$$\left.\begin{aligned}\bar{x}_1 &= \frac{1}{n_1}\sum_{i=1}^{n_1} x_{1i}\\ \bar{x}_2 &= \frac{1}{n_2}\sum_{j=1}^{n_2} x_{2j}\end{aligned}\right\}\ldots(*) \quad \text{and} \quad \left.\begin{aligned}\sigma_1^2 &= \frac{1}{n_1}\sum_{i=1}^{n_1}(x_{1i} - \bar{x}_1)^2\\ \sigma_2^2 &= \frac{1}{n_2}\sum_{j=1}^{n_2}(x_{2j} - \bar{x}_2)^2\end{aligned}\right\} \qquad \ldots(**)$$

The mean $\bar{x}$ of the combined series is given by

$$\bar{x} = \frac{1}{n_1 + n_2}\left[\sum_{i=1}^{n_1} x_{1i} + \sum_{j=1}^{n_2} x_{2j}\right] = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} \qquad [\text{From } (*)]$$

The variance $\sigma^2$ of the combined series is given by

$$\sigma^2 = \frac{1}{n_1 + n_2}\left[\sum_{i=1}^{n_1}(x_{1i} - \bar{x})^2 + \sum_{j=1}^{n_2}(x_{2j} - \bar{x})^2\right] \qquad \ldots(3{\cdot}10)$$

Now

$$\sum_{i=1}^{n_1}(x_{1i} - \bar{x})^2 = \sum_{i=1}^{n_1}(x_{1i} - \bar{x}_1 + \bar{x}_1 - \bar{x})^2$$

$$= \sum_{i=1}^{n_1}(x_{1i} - \bar{x}_1)^2 + n_1(\bar{x}_1 - \bar{x})^2 + 2(\bar{x}_1 - \bar{x})\sum_{i=1}^{n_1}(x_{1i} - \bar{x}_1). \qquad (3{\cdot}10a)$$

But $\sum_{i=1}^{n_1}(x_{1i} - \bar{x}_1) = 0$, being the algebraic sum of the deviations the values of first series from their mean. Hence from (3·10$a$), on using (**), we get

$$\sum_{i=1}^{n_1}(x_{1i} - \bar{x})^2 = n_1\sigma_1^2 + n_1(\bar{x}_1 - \bar{x})^2 = n_1\sigma_1^2 + n_1 d_1^2 \qquad \ldots(3{\cdot}10b)$$

where $d_1 = \bar{x}_1 - \bar{x}$.

Similarly, we get

$$\sum_{j=1}^{n_2} (x_{2j} - \bar{x})^2 = \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2 + \bar{x}_2 - \bar{x})^2$$

$$= \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)^2 + n_2 (\bar{x}_2 - \bar{x})^2 = n_2 \sigma_2^2 + n_2 d_2^2 \qquad ...(3\cdot10c)$$

where $\qquad d_2 = \bar{x}_2 - \bar{x}$.

Substituting from (3·10*b*) and (3·10*c*) in (3·10), we get the required formula

$$\sigma^2 = \frac{1}{n_1 + n_2} \left[ n_1 (\sigma_1^2 + d_1^2) + n_2 (\sigma_2^2 + d_2^2) \right]$$

This formula can be simplified still further. We have

$$d_1 = \bar{x}_1 - \bar{x} = \bar{x}_1 - \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2} = \frac{n_2 (\bar{x}_1 - \bar{x}_2)}{n_1 + n_2}$$

$$d_2 = \bar{x}_2 - \bar{x} = \bar{x}_2 - \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2} = \frac{n_1 (\bar{x}_2 - \bar{x}_1)}{n_1 + n_2}$$

Hence

$$\sigma^2 = \frac{1}{n_1 + n_2} \left[ n_1 \sigma_1^2 + n_2 \sigma_2^2 + \left\{ \frac{n_1 n_2^2 (\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2} + \frac{n_2 n_1^2 (\bar{x}_2 - \bar{x}_1)^2}{(n_1 + n_2)^2} \right\} \right]$$

$$= \frac{1}{n_1 + n_2} \left[ n_1 \sigma_1^2 + n_2 \sigma_2^2 + \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_1 - \bar{x}_2)^2 \right] \qquad ...(3\cdot11)$$

**Remark.** The formula (3·9) can be easily generalised to the case of more than two series. If $n_i$, $\bar{x}_i$ and $\sigma_i$, $i = 1, 2, ..., k$ are the sizes, means and standard deviations respectively of $k$-component series then the standard deviation $\sigma$ of the combined series of size $\sum_{i=1}^{k} n_i$ is given by

$$\sigma^2 = \frac{1}{n_1 + n_2 + \ldots + n_k} \left[ n_1 (\sigma_1^2 + d_1^2) + n_2 (\sigma_2^2 + d_2^2) + \ldots + n_k (\sigma_k^2 + d_k^2) \right] \qquad ...(3\cdot12)$$

where $\qquad d_i = \bar{x}_i - \bar{x}$ ; $i = 1, 2, ..., k$

and $\qquad \bar{x} = \dfrac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + \ldots + n_k \bar{x}_k}{n_1 + n_2 + \ldots + n_k}$

**Example 3·6.** *The first of the two samples has 100 items with mean 15 and standard deviation 3. If the whole group has 250 items with mean 15·6 and standard deviation $\sqrt{13\cdot44}$, find the standard deviation of the second group.*

**Solution.** Here we are given

$$n_1 = 100, \ \bar{x}_1 = 15 \ \text{and} \ \sigma_1 = 3$$

$$n = n_1 + n_2 = 250, \ \bar{x} = 15\cdot6, \ \text{and} \ \sigma = \sqrt{13\cdot44}$$

We want $\sigma_2$.

Obviously $\qquad n_2 = 250 - 100 = 150$. We have

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2} \Rightarrow 15{\cdot}6 = \frac{100 \times 15 + 150 \times \bar{x}_2}{250}$$

$\Rightarrow \quad 150\, \bar{x}_2 = 250 \times 15{\cdot}6 - 1500 = 2400$

$\therefore \quad \bar{x}_2 = \dfrac{2400}{150} = 16$

Hence $\quad d_1 = \bar{x}_1 - \bar{x} = 15 - 15{\cdot}6 = -0{\cdot}6$

and $\quad d_2 = \bar{x}_2 - \bar{x} = 16 - 15{\cdot}6 = 0{\cdot}4$

The variance $\sigma^2$ of the combined group is given by the formula :

$$(n_1 + n_2)\sigma^2 = n_1 (\sigma_1^2 + d_1^2) + n_2 (\sigma_2^2 + d_2^2)$$

$\Rightarrow \quad 250 \times 13{\cdot}44 = 100\,(9 + 0{\cdot}36) + 150\,(\sigma_2^2 + 0{\cdot}16)$

$\therefore \quad 150\, \sigma_2^2 = 250 \times 13{\cdot}44 - 100 \times 9{\cdot}36 - 150 \times 0{\cdot}16$

$= 3360 - 936 - 24 = 2400$

$\therefore \quad \sigma_2^2 = \dfrac{2400}{150} = 16$

Hence $\quad \sigma_2 = \sqrt{16} = 4$

**3·8. Co-efficient of Dispersion.** Whenever we want to compare the variability of the two series which differ widely in their averages or which are measured in different units, we do not merely calculate the measures of dispersion but we calculate the co-efficients of dispersion which are pure numbers independent of the units of measurement. The co-efficients of dispersion (C.D.) based on different measures of dispersion are as follows :

1. C.D. based upon range $= \dfrac{A - B}{A + B}$, where $A$ and $B$ are the greatest and the smallest items in the series.

2. Based upon quartile deviation :

$$\text{C.D.} = \frac{(Q_3 - Q_1)/2}{(Q_3 + Q_1)/2} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

3. Based upon mean deviation :

$$\text{C.D.} = \frac{\text{Mean deviation}}{\text{Average from which it is calculated}}$$

4. Based upon standard deviation :

$$\text{C.D.} = \frac{\text{S.D.}}{\text{Mean}} = \frac{\sigma}{\bar{x}}$$

**3·8·1. Co-efficient of Variation.** 100 times the co-efficient of desperison based upon standard deviation is called co-efficient of variation (C.V.),

$$\text{C.V.} = 100 \times \frac{\sigma}{\bar{x}} \qquad \ldots(3{\cdot}13)$$

According to Professor Karl Pearson who suggested this measure, C.V. *is the percentage variation in the mean, standard deviation being considered as the total variation in the mean.*

For comparing the variability of two series, we calculate the co-efficient of variations for each series. The series having greater C.V. is said to be more variable than the other and the series having lesser C.V. is said to be more consistent (or homogenous) than the other.

**Example 3·7.** *An analysis of monthly wages paid to the workers of two firms A and B belonging to the same industry gives the following results :*

| | *Firm A* | *Firm B* |
|---|---|---|
| *Number of workers* | *500* | *600* |
| *Average monthly wage* | *Rs. 186·00* | *Rs. 175·00* |
| *Variance of distribution of wages* | *81* | *100* |

*(i) Which firm, A or B, has a larger wage bill ?*

*(ii) In which firm, A or B, is there greater variability in individual wages?*

*(iii) Calculate (a) the average monthly wage, and (b) the variance of the distribution of wages, of all the workers in the firms A and B taken together.*

**Solution.**

*(i)* **Firm A :**

No. of wage-earners (say) $n_1 = 500$
Average monthly wages (say) $\bar{x}_1$ = Rs.186

$$\text{Average monthly wage} = \frac{\text{Total wages paid}}{\text{No. of workers}}$$

Hence total wages paid to the workers = $n_1\bar{x}_1 = 500 \times 186$ = Rs. 93,000

**Firm B**

No. of wage-earners (say) $n_2 = 600$
Average monthly wages (say) $\bar{x}_2$ = Rs.175

$\therefore$ Total wages paid to the workers = $n_2\bar{x}_2 = 600 \times 175$ = Rs. 1,05,000

Thus we see that the firm *B* has larger wage bill.

*(ii)* Variance of distribution of wages in firm *A* (say) $\sigma_1^2 = 81$
Variance of distribution of wages in firm *B* (say) $\sigma_2^2 = 100$

$$\text{C.V. of distribution of wages for firm } A = 100 \times \frac{\sigma_1}{\bar{x}_1} = \frac{100 \times 9}{186} = 4{\cdot}84$$

$$\text{C.V. of distribution of wages for firm } B = 100 \times \frac{\sigma_2}{\bar{x}_2} = \frac{100 \times 10}{175} = 5{\cdot}71$$

Since C.V. for firm *B* is greater than C.V. for firm *A*, firm *B* has greater variability in individual wages.

*(iii)* *(a)* The average monthly wages (say) $\bar{x}$, of all the workers in the two firms A and *B* taken together is given by

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} = \frac{500 \times 186 + 600 \times 175}{500 + 600} = \frac{198000}{1100} = \text{R}\ \ 180$$

*(b)* The combined variance $\sigma^2$ is given by the formula:

$$\sigma^2 = \frac{1}{n_1 + n_2}\left[ n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)\right]$$

where $d_1 = \bar{x}_1 - \bar{x}$ and $d_2 = \bar{x}_2 - \bar{x}$

Here $d_1 = 186 - 180 = 6$ and $d_2 = 175 - 180 = -5$

Hence $\sigma^2 = \dfrac{500\,(81 + 36) + 600\,(100 + 25)}{500 + 600} = \dfrac{133500}{1100} = 121{\cdot}36$

## EXERCISE 3 (a)

1. (a) Explain with suitable examples the term 'dispersion.' State the relative and absolute measures of dispersion and describe the merits and demerits of standard deviation.

*(b)* Explain the main difference between mean deviation and standard deviation. Show that standard deviation is independent of change of origin and scale.

*(c)* Distinguish between absolute and retative measurés of dispersion.

2. *(a)* Explain the graphical method of obtaining median and quartile deviation. (Calicut Univ.B.Sc, .April 1989)

*(b)* Compute quartile deviation graphically for the following data :

| Marks : | 20 – 30 | 30 – 40 | 40 – 50 | 50 – 60 | 60 – 70 | 70 & over |
|---|---|---|---|---|---|---|
| Number of students : | 5 | 20 | 14 | 10 | 8 | 5 |

3. *(a)* Show that for raw data mean deviation is minimum when measured from the median.

*(b)* Compute a suitable measure of dispersion for the following grouped frequency distribution giving reasons :

| *Classes* | *Frequency* |
|---|---|
| Less than 20 | 30 |
| 20 – 30 | 20 |
| 30 – 40 | 15 |
| 40 – 50 | 10 |
| 50 – 60 | 5 |

(c) Age distribution of hundred life insurance policyholders is as follows:

| *Age as on nearest birthday* | *Number* |
|---|---|
| 17 – 19·5 | 9 |
| 20 – 25·5 | 16 |
| 26 – 35·5 | 12 |
| 36 – 40·5 | 26 |
| 41 – 50·5 | 14 |
| 51 – 55·5 | 12 |
| 56 – 60·5 | 6 |
| 61 – 70·5 | 5 |

Calculate mean deviation from median age.

Ans. Median = 38·25, M.D.=10·605

4. Prove that the mean deviation about the mean $\bar{x}$ of the variate x, the frequency of whose ith size $x_i$ is $f_i$ is given by

$$\frac{2}{N}\left[\bar{x}\sum_{x_i<\bar{x}} f_i - \sum_{x_i<\bar{x}} f_i\, x_i\right].$$

**Hint.** Mean deviation about mean

$$= \frac{1}{N}\left[\sum_{x_i<\bar{x}} f_i(\bar{x}-x_i) + \sum_{x_i>\bar{x}} f_i(x_i-\bar{x})\right]$$

$$= \frac{1}{N}\left[-\sum_{x_i<\bar{x}} f_i(x_i-\bar{x}) + \sum_{x_i>\bar{x}} f_i(x_i-\bar{x})\right]$$

Since $\Sigma f_i(x_i-\bar{x}) = 0,$

$$\sum_{x_i>\bar{x}} f_i(x_i-\bar{x}) + \sum_{x_i<\bar{x}} f_i(x_i-\bar{x}) = 0$$

$$\therefore\ M.D. = \frac{1}{N}\left(-\sum_{x_i<\bar{x}} f_i(x_i-\bar{x}) - \sum_{x_i<\bar{x}} f_i(x_i-\bar{x})\right) = -\frac{2}{N}\left(\sum_{x_i<\bar{x}} f_i(x_i-\bar{x})\right)$$

5. What is standard deviation ? Explain its superiority over other measures of dispersion.

6. Calculate the mean and standard deviation of the following distribution:

| | | | | |
|---|---|---|---|---|
| x : | 2.5 — 7.5 | 7.5 — 12.5 | 12.5 — 17.5 | 17.5 — 22.5 |
| f : | 12 | 28 | 65 | 121 |

| | | | | | |
|---|---|---|---|---|---|
| x : | 22.5 — 27.5 | 27.5 — 32.5 | 32.5 — 37.5 | 37.5 — 42.5 | 42 5 — 47.5 |
| f : | 175 | 198 | 176 | 120 | 66 |

| | | | |
|---|---|---|---|
| x : | 47.5 — 52.5 | 52.5 — 57.5 | 57.5 — 62.5 |
| f : | 27 | 9 | 3 |

**Ans.** Mean = 30.005, Standard Deviation = 0.01

7. Explain clearly the ideas implied in using arbitrary working orgin, and scale for the calculation of the arithmetic mean and standard deviation of a frequency distribution. The values of the arithmetic mean and standard deviation of the following frequency distribution of a continuous variable derived from the analysis in the above manner are40.604 lb. and 7.92 lb. respectively.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $x$ : | –3 | –2 | –1 | 0 | 1 | 2 | 3 | 4 | Total |
| $f$ : | 3 | 15 | 45 | 57 | 50 | 36 | 25 | 9 | 240 |

Determine the actual class intervals.

8. *(a)*. The arithmetic mean and variance of a set of 10 figures are known to be 17 and 33 respectively. Of the 10 figures, one figure (*i.e.*, 26) was subsequently found inacurate, and was weeded out. What is the resulting *(a)* arithmetic mean and *(b)* standard deviation. (M.S. Baroda U. B.Sc. 1993)

*(b)* The mean and standard deviation of 20 items is found to be 10 and 2 respectively. At the time of checking it was found that one item 8 was incorrect. Calculate the mean and standard deviation if

*(i)* the wrong item is omitted, and
*(ii)* it is replaced by 12.

(c) For a frequency distribution of marks in Statistics of 200 candidates (grouped in intervals 0–5, 5–10,..., etc.), the mean and standard deviation were found to be 40 and 15 respectively. Later it was discovered that the score 43 was misread as 53 in obtaining the frequency distribution. Find the corrected mean and standard deviation corresponding to the corrected frequency distribution.

**Ans.** Mean = 39·95, S.D. = 14·974.

**9.** *(a)* Complete a table showing the frequencies with which words of different numbers of letters occur in the extract reproduced below (omitting punctuation marks) treating as the variable the number of letters in each word, and obtain the mean, median and co-efficient of variation of the distribution :

"Her eyes were blue : blue as autumn distance-blue as the blue we see, between the retreating mouldings of hills and woody slopes on a sunny September morning : a misty and shady blue, that had no beginning or surface, and was looked into rather than at. "

**Ans.** Mean = 4·35, Median = 4, $\sigma$ = 2·23 and C.V. = 51·26

*(b)* Treating the number of letters in each word in the following passage as the variable x, prepare the frequency distribution table and obtain its mean, median, mode and variance.

"The reliability of data must always be examined before any attempt is made to base conclusions upon them. This is true of all data, but particularly so of numerical data, which do not carry their quality written large on them. It is a waste of time to apply the refined theoretical methods of Statistics to data which are suspect from the beginning. "

**Ans.** Mean = 4·565, Median = 4, Mode = 3, S.D. = 2·673.

**10.** The mean of 5 observations is 4·4 and variance is 8·24. It three of the five observation are 1, 2 and 6, find the other two.

**11.** *(a)* Scores of two golfers for 24 rounds were as follows :

Golfer A : 74, 75, 78, 72, 77, 79, 78, 81, 76, 72, 72, 77, 74, 70, 78, 79, 80, 81, 74, 80, 75, 71, 73.

Golfer B : 86, 84, 80, 88, 89, 85, 86, 82, 82, 79, 86, 80, 82, 76, 86, 89, 87, 83, 80, 88, 86, 81, 81, 87

Find which golfer may be considered to be a more consistent player ?

**Ans.** Golfer B is more consistent player.

*(b)* The sum and sum of squares corresponding to length $X$ (in cms.) and weight $Y$ (in gms.) of 50 tapioca tubers are given below :

$\Sigma X = 212, \quad \Sigma X^2 = 902 \cdot 8$

$\Sigma Y = 261, \quad \Sigma Y^2 = 1457 \cdot 6$

Which is more varying, the length or weight.

**12.** *(a)* Lives of two models of refrigerators turned in for new models in a recent survey are

| *Life (No. of years)* | *Model A* | *Model B* |
|---|---|---|
| 0 – 2 | 5 | 2 |
| 2 – 4 | 16 | 7 |

| | | |
|---|---|---|
| 4 – 6 | 13 | 12 |
| 6 – 8 | 7 | 19 |
| 8 – 10 | 5 | 9 |
| 10 – 12 | 4 | 1 |

What is the average life of each model of these refrigerators ? Which model shows more uniformity ?

**Ans.** C.V. (Model A)=54·9%, C.V. (Model B)=3·62%

(*b*) Goals scored by two teams A and B in a football season were as follows :

| *No. of goals scored in a match* | *No. of matches* A | B |
|---|---|---|
| 0 | 27 | 17 |
| 1 | 9 | 9 |
| 2 | 8 | 6 |
| 3 | 5 | 5 |
| 4 | 4 | 3 |

**(Sri Venketeswara U. B.Sc. Sept. 1992)**

Find out which team is more consistent.

**Ans.** Team $A$ : C.V.= 122·0, Team $B$ : C.V. = 108·3 .

(c) An analysis of monthly wages paid to the workers in two firms, A and B belonging to the same industry, gave the following results :

| | *Firm A* | *Firm B* |
|---|---|---|
| No. of wage-earners | 986 | 548 |
| Average monthly wages | Rs. 52·5 | Rs. 47·5 |
| Variance of distribution of wages | 100 | 121 |

(*i*) Which firm, $A$ or $B$, pays out larger amount as monthly wages ?

(*ii*) In which firm $A$ or $B$, is there greater variability in individual wages?

(*iii*) What are the measures of average monthly wages and the variability in individual wages, of all the workers in the two firms, $A$ and $B$ taken together.

**Ans.** (*i*) Firm $B$ pays a larger amount as monthly wages.
(*ii*) There is greater variability in individual wages in firm $B$.
(*iii*) Combined arithmetic mean = Rs.49·87.
Combined standard deviation = Rs.10·82.

14. (*a*) The following data give the arithmetic averages and standard deviations of three sub-groups. Calculate the arithmetic average and standard deviation of the whole group.

| *Sub-group* | *No. of men* | *Average wages (Rs.)* | *Standard deviation (Rs.)* |
|---|---|---|---|
| A | 50 | 61·0 | 8·0 |
| B | 100 | 70·0 | 9·0 |
| C | 120 | 80·5 | 10·0 |

**Ans.** Combined Mean = 73, Combined S.D.=11·9

(*b*) Find the missing information from the following data :

| | Group I | Group II | Group III | Combined |
|---|---|---|---|---|
| Number | 50 | ? | 90 | 200 |
| Standard Deviation | 6 | 7 | ? | 7·746 |
| Mean | 113 | ? | 115 | 116 |

**Ans.** $n_2 = 60$, $\bar{x}_2 = 120$ and $\sigma_3 = 8$

**15.** A collar manufacturer is considering the production of a new style collar to attract young men. The following statistics of neck circumference are avail based on the measurement of a typical group of students :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Mid-value in inches | : 12·5 | 13·0 | 13·5 | 14·0 | 14·5 | 15·0 | 15·5 | 16·0 |
| No. of students | : 4 | 19 | 30 | 63 | 66 | 29 | 18 | 1 |

Compute the mean and standard deviation and use the criterion $\bar{x} \pm$ obtain the largest and smallest size of collar he should make in order to meet needs of practically all his customers bearing in mind that the collars are worn on average 3/4 inch larger than neck size. **(Nagpur Univ. B.Sc., 1992)**

**Ans.** Mean = 14·232, S.D.=0·72, largest size = 17·14", smallest size = 12·83"

**16.** *(a)* A frequency distribution is divided into two parts. The mean and standard deviation of the first part are $m_1$ and $s_1$ and those of the second part are $m_2$ and $s_2$ respectively. Obtain the mean and standard deviation for the combined distribution. **[Delhi Univ. B.Sc.(Stat.Hons.), 1986]**

*(b)* The means of two samples of size 50 and 100 respectively are 54·1 and 50·3 and the standard deviations are 8 and 7. Obtain the mean and standard deviation of the sample of size 150 obtained by combining the two samples.

**Ans.** Combined mean = 51·57. Combined S.D. = 7·5 approx.

*(c)* A distribution consists of three components with frequencies 200, 250 and 300 having means 25, 10 and 15 and standard deviations 3, 4 and 5 respectively.

Show that the mean of the combined group is 16 and its standard deviation 7.2 approximately. **(Bangalore Univ. B.Sc. 1992)**

**17.** In a certain test for which the pass marks is 30, the distribution of marks of passing candidates classified by sex (boys and girls) were as given below :

| *Marks* | *Frequency* | |
|---|---|---|
| | *Boys* | *Girls* |
| 30-34 | 5 | 15 |
| 35-39 | 10 | 20 |
| 40-44 | 15 | 30 |
| 45-49 | 30 | 20 |
| 50-54 | 5 | 5 |
| 55-59 | 5 | - |
| Total | 70 | 90 |

The overall means and standard deviation of marks for boys including the 30 failed were 38 and 10. The corresponding figures for girls including the 10 failed were 35 and 9.

*(i)* Find the mean and standard deviation of marks obtained by the 30 boys who failed in the test.

*(ii)* The moderation committee argued that percentage of passes among girls is higher because the girls are very studious and if the intention is to pass those who are really intelligent, a higher pass marks should be used for girls. Wiothout quetioning the propriety of this argument, suggest what the pass mark should be which would allow only 70% of the girls to pass.

*(iii)* The prize committee decided to award prizes to the best 40 candidates (irrespective of sex) judged on the basis of marks obtained in the test. Estimate the number of girls who would receive prizes.

**Ans.** (i) $\bar{x} = 22{\cdot}83$, $\sigma_2 = 8{\cdot}27$ (ii) 39 (iii) 15

**18.** Find the mean and variance of first n-natural numbers.

**(Agra Univ. B.Sc., 1993)**

**Ans.** $\bar{x} = \dfrac{n+1}{2}$, $\sigma_2 = \dfrac{n^2-1}{12}$

**19.** In a frequency distribution, the $n$ intervals are 0 to 1, 1 to 2, ..., $(n-1)$ to $n$ with equal frequencies. Find the mean deviation and variance.

**20.** If the mean and standard deviation of a variable $x$ are $m$ and $\sigma$ respectively, obtain the mean and standard deviation of $(ax+b)/c$, where $a$, $b$ and $c$ are constants.

**Ans.** $\bar{u} = \dfrac{1}{c}(a\bar{x}+b)$, $\sigma_u = \left|\dfrac{a}{c}\right| \sigma$

**21.** In a series of measurements we obtain $m_1$ values of magnitude $x_1$, $m_2$ values of magnitude $x_2$, and so on. If $\bar{x}$ is the mean value of all the measurements, prove that the standard deviation is

$$\sqrt{\frac{\Sigma m_r (k - x_r)^2}{\Sigma m_r} - \delta^2}$$

where $\bar{x} = k + \delta$ and $k$ is any constant. **Delhi Univ. B.Sc. (Stat. Hons.), 1992**

**22.** (a) Show that in a discrete series if deviations are small compared with mean M so that $(x/M)^2$ and higher powers of $(x/M)$ are neglected, prove that

*(i)* $MH = G^2$ *(II)* $M - 2G + H = 0$,

where $G$ is geometric mean and $H$ is harmonic mean.

(b) The mean and standard deviation of a variable x are m and $\sigma$ respectively. If the deviations are small compared with the value of the mean, show that

(i) Mean $(\sqrt{x}) = \sqrt{m}\left(1 - \dfrac{\sigma^2}{8m^2}\right)$

(ii) Mean $\left(\dfrac{1}{\sqrt{x}}\right) = \dfrac{1}{\sqrt{m}}\left(1 + \dfrac{3\sigma^2}{8m^2}\right)$ approximately.

**(M.S. Baroda U. B.Sc. 1993)**

(c) If the deviation $X_i = x_i - M$ is very small in comparison with mean $M$ and $(X_i/M)^2$ and higher powers of $(X_i/M)$ are neglected, prove that

$$V = \sqrt{\frac{2(M-G)}{M}}$$

where $G$ is the geometric mean of the values $x_1, x_2, \ldots, x_n$ and $V$ is the coefficient of dispersion $(\sigma/M)$. **(Lucknow Univ. B.Sc., 1993)**

**23.** From a sample of observations the arithmetic mean and variance are calculated. It is then found that one of the values, $x_1$, is in error and should be replaced by $x_1'$. Show that the adjustment to the variance to correct this error is

$$\frac{1}{n}(x'_1 - x_1)\left(x'_1 + x_1 - \frac{x_1' - x_1 + 2T}{n}\right)$$

where $T$ is the total of the original results.

**(Meerut Univ. B.Sc., 1992; Delhi Univ. B.Sc. (Stat. Hons.), 1989, 1985]**

**Hint.** $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$

$$= \frac{1}{n}\left(x_1^2 + x_2^2 + \ldots + x_n^2\right) - \frac{T^2}{n^2}$$

where $T = x_1 + x_2 + \ldots + x_n$.

Let $\sigma_1^2$ be the corrected variance. Then

$$\sigma_1^2 = \frac{1}{n}\left\{x_1'^2 + x_2^2 + \ldots + x_n^2\right\} - \left\{\frac{T - x_1 + x_1'}{n}\right\}^2$$

Adjustment to the variance to correct the error is :

$$\sigma_1^2 - \sigma^2 = \frac{1}{n}\left\{x_1'^2 - x_1^2\right\} - \frac{1}{n^2}\left\{\left(T - x_1 + x_1'\right)^2 - T^2\right\}$$

$$= \frac{1}{n}\left\{x_1' + x_1\right\}\left\{x_1' - x_1\right\} - \frac{1}{n^2}\left\{\left(x_1' - x_1\right) \times (2T - x_1 + x_1')\right\}$$

**24.** Show that, if the variable takes the values $0, 1, 2, \ldots, n$ with frequencies proportional to the binomial coefficients ${}^nC_0, {}^nC_1, {}^nC_2, \ldots, {}^nC_n$ respectively then the mean of the distribution is $(n/2)$, the mean square deviation about $x = 0$ is $n(n+1)/4$ and the variance is $n/4$

**[Delhi Univ. B.Sc. (Stat. Hons.), 1991]**

**Hint.** $N = \sum f = {}^nC_0 + {}^nC_1 + {}^nC_2 + \ldots + {}^nC_n = (1+1)^n = 2^n$

$$\sum fx = 0.{}^nC_0 + 1.{}^nC_1 + 2.{}^nC_2 + 3.{}^nC_3 + \ldots + n.{}^nC_n$$

$$= n\left\{1 + (n-1) + \frac{(n-1)(n-2)}{2!} + \ldots + 1\right\}$$

$$= n(1+1)^{n-1} = n \cdot 2^{(n-1)}$$

Hence mean $(\bar{x}) = \frac{1}{N}\sum fx = \frac{n \cdot 2^{(n-1)}}{2^n} = \frac{n}{2}$

The mean square deviation $s^2$, (say), about the point $x = 0$ is given by

$$s^2 = \frac{1}{N}\sum fx^2 = \frac{1}{2^n}\left[1^2.{}^nC_1 + 2^2.{}^nC_2 + 3^2.{}^nC_3 + \ldots + n^2.{}^nC_n\right]$$

$$= \frac{n}{2^n}\left[1 + 2(n-1) + \frac{3}{2}(n-1)(n-2) + \ldots + n\right]$$

$$= \frac{n}{2^n}\left(\left\{1 + (n-1) + \frac{(n-1)(n-2)}{2!} + \ldots + 1\right\} + \left\{(n-1) + (n-1)(n-2) + \ldots + (n-1)\right\}\right)$$

$$= \frac{n}{2^n}\left[\left({}^{n-1}C_0 + {}^{n-1}C_1 + {}^{n-1}C_2 + \ldots + {}^{n-1}C_{n-1}\right) + \left\{(n-1)\left({}^{n-2}C_0 + {}^{n-2}C_1 + \ldots + {}^{n-2}C_{n-2}\right)\right\}\right]$$

$$= \frac{n}{2^n}\left[(1+1)^{n-1} + (n-1)(1+1)^{n-2}\right] = \frac{n(n+1)}{4}$$

$$\therefore \qquad \sigma^2 = \frac{n(n+1)}{4} - \frac{n^2}{4} = \frac{n}{4}.$$

**25.** (*a*) Let $r$ be the range and $s$ be the standard deviation of a set of observations $x_1, x_2, \ldots, x_n$; then prove by general reasoning or otherwise that $s \leq r$.

**Hint.** Since $x_i - \bar{x} \leq r$, $i = 1, 2, \ldots, n$, we have

$$s^2 = \frac{1}{N}\sum_{i=1}^{n} f_i(x_i - \bar{x})^2 \leq \frac{1}{N}\sum_{i=1}^{n} f_i\,(r^2)$$

$$\Rightarrow \qquad s^2 \leq r^2 \frac{1}{N}\sum_{i=1}^{n} f_i = r^2 \quad \Rightarrow \quad s \leq r$$

(*b*) Let $r$ be the range and

$$S = \left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{\frac{1}{2}}$$

be the standard deviation of a set of observations $x_1, x_2, \ldots, x_n$, then prove that

$$S \leq r\left(\frac{n}{n-1}\right)^{\frac{1}{2}}$$ **[Punjab Univ. B.Sc (Stat. Hons.), 1993]**

**3·9. Moments.** The $r$th moment of a variable $x$ about any point $x = A$, usually denoted by $\mu_r'$ is given by

$$\mu_r' = \frac{1}{N}\sum_i f_i(x_i - A)^r, \quad \sum_i f_i = N \qquad \ldots(3\cdot14)$$

$$= \frac{1}{N}\sum_i f_i d_i^r, \qquad \ldots(3\cdot14a)$$

where $d_i = x_i - A$.

The $r$th moment of a variable about the mean $\bar{x}$, usually denoted by $\mu_r$ is given by

$$\mu_r = \frac{1}{N}\sum_i f_i(x_i - \bar{x})^r = \frac{1}{N}\sum_i f_i z_i^r \qquad \ldots(3\cdot15)$$

where $\qquad z_i = x_i - \bar{x}$.

In particular

$$\mu_0 = \frac{1}{N}\sum_i f_i(x_i - \bar{x})^0 = \frac{1}{N}\sum_i f_i = 1$$

and $\mu_1 = \dfrac{1}{N}\sum_i f_i(x_i - \bar{x}) = 0$, being the algebraic sum of deviations from the mean. Also

$$\mu_2 = \frac{1}{N}\sum_i f_i(x_i - \bar{x})^2 = \sigma^2 \qquad ...(3·16)$$

These results, viz., $\mu_0 = 1$, $\mu_1 = 0$, and $\mu_2 = \sigma^2$, are of fundamental importance and should be committed to memory.

We know that if $d_i = x_i - A$, then

$$\bar{x} = A + \frac{1}{N}\sum_i f_i d_i = A + \mu_1' \qquad ...(3·17)$$

**3·9·1. Relation between moments about mean in terms of moments about any point and *vice versa.***

We have

$$\mu_r = \frac{1}{N}\sum_i f_i(x_i - \bar{x})^r = \frac{1}{N}\sum_i f_i(x_i - A + A - \bar{x})^r$$

$$= \frac{1}{N}\sum_i f_i(d_i + A - \bar{x})^r, \text{ where } d_i = x_i - A$$

Using (3·17), we get

$$\mu_r = \frac{1}{N}\sum_i f_i(d_i - \mu_1')^r$$

$$= \frac{1}{N}\Sigma f_i\left[d_i^r - {}^rC_1 d_i^{r-1}\mu_1' + {}^rC_2 d_i^{r-2}\mu_1'^2 - {}^rC_3 d_i^{r-3}\mu_1'^3 + \ldots + (-1)^r\mu_1'^r\right]$$

...(3·18)

$= \mu_r' - {}^rC_1\mu_{r-1}'\mu_1' + {}^rC_2\mu_{r-2}'\mu_1'^2 - \ldots + (-1)^r\mu_1'^r$ [On using (3·14a)]

In particular, on putting r = 2, 3 and 4 in (3·18), we get

$$\mu_2 = \mu_2' - \mu_1'^2$$
$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 \qquad ...(3.19)$$
$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$$

Conversely,

$$\mu_r' = \frac{1}{N}\sum_i f_i(x_i - A)^r = \frac{1}{N}\sum_i f_i(x_i - \bar{x} + \bar{x} - A)^r$$

$$= \frac{1}{N}\sum_i f_i\,(z_i + \mu_1')^r$$

where $x_i - \bar{x} = z_i$ and $\bar{x} = A + \mu_1'$

Thus $\mu_r' = \frac{1}{N} \sum_i f_i \left( z_i^r + {}^rC_1 z_i^{r-1} \mu_1' + {}^rC_2 z_i^{r-2} \mu_1'^2 + \ldots + \mu_1'^r \right)$

$= \mu_r + {}^rC_1 \mu_{r-1} \mu_1' + {}^rC_2 \mu_{r-2} \mu_1'^2 + \ldots + \mu_1'^r$. [From (3·15)]

In particular, putting $r = 2, 3$ and $4$ and noting that $\mu_1 = 0$, we get

$$\mu_2' = \mu_2 + \mu_1'^2$$
$$\mu_3' = \mu_3 + 3\mu_2 \mu_1' + \mu_1'^3 \qquad \ldots (3·20)$$
$$\mu_4' = \mu_4 + 4\mu_3\mu_1' + 6\mu_2 \mu_1'^2 + \mu_1'^4$$

These formulae enable us to find the moments about any point, once the mean and the moments about mean are known.

### 3.9.2 Effect of Change of Origin and Scale on Moments.

Let $u = \frac{x - A}{h}$, so that $x = A + hu$, $\bar{x} = A + h\bar{u}$ and $x - \bar{x} = h(u - \bar{u})$

Thus, $r$th moment of $x$ about any point $x = A$ is given by

$$\mu_r' = \frac{1}{N} \sum_i f_i (x_i - A)^r = \frac{1}{N} \sum_i f_i \ (hu_i)^r = h^r \cdot \frac{1}{N} \cdot \sum_i f_i u_i^r$$

And the $r$th moment of $x$ about mean is

$$\mu_r = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^r = \frac{1}{N} \sum_i f_i \,[h(u_i - \bar{u})]^r$$
$$= h^r \frac{1}{N} \sum_i f_i (u_i - \bar{u})^r$$

Thus the $r$th moment of the variable $x$ about mean is $h^r$ times the $r$th moment of the variable $u$ about its mean.

**3·9·3. Sheppard's Corrections for Moments.** In case of grouped frequency distribution, while calculating moments we assume that the frequencies are concentrated at the middle point of the class intervals. If the distribution is symmetrical or slightly symmetrical and the class intervals are not greater than one-twentieth of the range, this assumption is very nearly true. But since the assumption is not in general true, some error, called the 'grouping error', creeps into the calculation of the moments. W.F. Sheppard proved that if

(*i*) the frequency distribution is continuous, and

(*ii*) the frequency tapers off to zero in both directions,

the effect due to grouping at the mid-point of the intervals can be corrected by the following formulae, known as Sheppard's corrections :

$$\mu_2 \text{ (corrected)} = \mu_2 - \frac{h^2}{12} \qquad \ldots (3·21)$$

$$\mu_3 \text{ (corrected)} = \mu_3$$

$$\mu_4 \text{ (corrected)} = \mu_4 - \frac{1}{2} h^2 \mu_2 + \frac{7}{240} h^4$$

where $h$ is the width of the class interval.

**3·9·4. Charlier's Checks.** The following identities

$\sum f(x+1) = \sum fx + N$; $\sum f(x+1)^2 = \sum fx^2 + 2\sum fx + N$

$\sum f(x+1)^3 = \sum fx^3 + 3\sum fx^2 + 3\sum fx + N$

$\sum f(x+1)^4 = \sum fx^4 + 4\sum fx^3 + 6\sum fx^2 + 4\sum fx + N,$

are often used in checking the accuracy in the calculation of first four moments and are known as Charlier's Checks.

**3·10. Pearson's β and γ Coefficients.** Karl Pearson defined the following four coefficients, based upon the first four moments about mean :

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \quad \gamma_1 = +\sqrt{\beta_1} \text{ and } \beta_2 = \frac{\mu_4}{\mu_2^2}, \quad \gamma_2 = \beta_2 - 3 \qquad \dots (3\cdot22)$$

It may be pointed out that these coefficients are pure numbers independent of units of measurement. The practical utility of these coefficients is discussed in § 3·13 and § 3·14.

**Remark.** Sometimes, another coefficient based on moments, *viz*, Alpha ($\alpha$) coefficient is used. Alpha coefficients are defined as :

$$\alpha_1 = \frac{\mu_1}{\sigma} = 0, \quad \alpha_2 = \frac{\mu_2}{\sigma^2} = 1, \quad \alpha_3 = \frac{\mu_3}{\sigma^3} = \sqrt{\beta_1} = \gamma_1; \quad \alpha_4 = \frac{\mu_4}{\sigma^4} = \beta_2$$

**3·11. Factorial Moments.** Factorial moment of order $r$ about the origin of the frequency distribution $x_i | f_i$, $(i = 1, 2, \dots n)$, is defined as

$$\mu_{(r)}' = \frac{1}{N} \sum_{i=1}^{n} f_i x_i^{(r)} \qquad \dots (3\cdot23)$$

where $x^{(r)} = x(x-1)(x-2)\dots(x-r+1)$ and $N = \sum_{i=1}^{n} f_i$

Thus the factorial moment of order $r$ about any point $x = a$ is given by

$$\mu_{(r)}' = \frac{1}{N} \sum_i f_i \, (x_i - a)^{(r)} \qquad \dots (3\cdot24)$$

where $(x-a)^{(r)} = (x-a)(x-a-1)\dots(x-a-r+1)$

In particular from (3·23), we have

$$\mu_{(1)}' = \frac{1}{N} \sum_i f_i x_i = \mu_1' \text{ (about origin)} = \text{Mean } (\bar{x})$$

$$\mu_{(2)}' = \frac{1}{N} \sum_i f_i x_i^{(2)} = \frac{1}{N} \sum_i f_i x_i (x_i - 1)$$

$$= \frac{1}{N} \sum_i f_i x_i^2 - \frac{1}{N} \sum_i f_i x_i = \mu_2' - \mu_1'$$

$$\mu_{(3)}' = \frac{1}{N}\sum_i f_i x_i^{(3)} = \frac{1}{N}\sum_i f_i x_i (x_i - 1)(x_i - 2)$$

$$= \frac{1}{N}\sum_i f_i x_i^3 - 3\,\frac{1}{N}\sum_i f_i x_i^2 + 2\,\frac{1}{N}\cdot\sum_i f_i x_i$$

$$= \mu_3' - 3\mu_2' + 2\mu_1'$$

$$\mu_{(4)}' = \frac{1}{N}\sum_i f_i x_i^{(4)} = \frac{1}{N}\sum_i f_i x_i (x_i - 1)(x_i - 2)(x_i - 3)$$

$$= \frac{1}{N}\sum_i f_i x_i (x_i^3 - 6x_i^2 + 11x_i - 6)$$

$$= \frac{1}{N}\sum_i f_i x_i^4 - 6.\,\frac{1}{N}\sum f_i x_i^3 + 11.\,\frac{1}{N}\sum f_i x_i^2 - 6.\,\frac{1}{N}\sum f_i x_i$$

$$= \mu_4' - 6\mu_3' + 11\mu_2' - 6\mu_1'$$

Conversely, we will get

$$\mu_1' = \mu_{(1)}'$$

$$\mu_2' = \mu_{(2)}' + \mu_{(1)}'$$

$$\mu_3' = \mu_{(3)}' + 3\mu_{(2)}' + \mu_{(1)}' \qquad \ldots (3\cdot25)$$

$$\mu_4' = \mu_{(4)}' + 6\mu_{(3)}' + 7\mu_{(2)}' + \mu_{(1)}'$$

**3·12. Absolute Moments.** For the frequency distribution $x_i / f_i$ $i = 1, 2, \ldots n$, the $r$th absolute moment of the variable about the origin is given by

$$\frac{1}{N}\sum_{i=1}^{n} f_i \,|\, x_i^r \,|,\; N = \sum f_i \qquad \ldots (3\cdot 26)$$

where $|\, x_i^r \,|$ represents the absolute or modulus value of $x_i^r$.

The $r$th absolute moment of the variable about the mean $\bar{x}$ is given by

$$\frac{1}{N}\sum_{i=1}^{n} f_i \,|\, x_i - \bar{x} \,|^r \qquad \ldots (3\cdot26a)$$

**Example 3·8.** *The first four moments of a distribution about the value 4 of the variable are* $-1{\cdot}5$, $17$, $-30$ *and* $108$. *Find the moments about mean,* $\beta_1$ *and* $\beta_2$.

*Find also the moments about (i) the origin, and (ii) the point* $x = 2$.

**Solution.** In the usual notations, we are given $A = 4$ and

$$\mu_1' = -1{\cdot}5,\; \mu_2' = 17,\; \mu_3' = -30 \text{ and } \mu_4' = 108.$$

*Moments about mean :* $\mu_1 = 0$

$$\mu_2 = \mu_2' - \mu_1'^2 = 17 - (-1{\cdot}5)^2 = 17 - 2{\cdot}25 = 14{\cdot}75$$

$$\mu_3 = \mu_3' - 3\mu_2'\,\mu_1' + 2\mu_1'^3$$

$$= -30 - 3 \times (17) \times (-1{\cdot}5) + 2\,(-1{\cdot}5)^3$$

$$= -30 + 76{\cdot}5 - 6{\cdot}75 = 39{\cdot}75$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$$
$$= 108 - 4(-30)(-1{\cdot}5) + 6(17)(-1{\cdot}5)^2 - 3(-1{\cdot}5)^4$$
$$= 108 - 180 + 229{\cdot}5 - 15{\cdot}1875 = 142{\cdot}3125$$

Hence $$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(39{\cdot}75)^2}{(14{\cdot}75)^3} = 0{\cdot}4924$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{(142{\cdot}3125)}{(14{\cdot}75)^2} = 0{\cdot}6541$$

Also $$\bar{x} = A + \mu_1' = 4 + (-1{\cdot}5) = 2{\cdot}5$$

*Moments about origin.* We have

$$\bar{x} = 2{\cdot}5,\ \mu_2 = 14{\cdot}75,\ \mu_3 = 39{\cdot}75 \text{ and } \mu_1 = 142{\cdot}31 \text{ (approx).}$$

We know $\bar{x} = A + \mu_1'$, where $\mu_1'$ is the first moment about the point $x = A$. Taking $A = 0$, we get the first moment about origin as $\mu_1' =$ mean $= 2{\cdot}5$.

Using (3·20), we get

$$\mu_2' = \mu_2 + \mu_1'^2 = 14{\cdot}75 + (2{\cdot}5)^2 = 14{\cdot}75 + 6{\cdot}25 = 21$$
$$\mu_3' = \mu_3 + 3\mu_2\mu_1' + \mu_1'^3 = 39{\cdot}75 + 3(14{\cdot}75)(2{\cdot}5) + (2{\cdot}5)^3$$
$$= 39{\cdot}75 + 110{\cdot}625 + 15{\cdot}625 = 166$$
$$\mu_4' = \mu_4 + 4\mu_3\mu_1' + 6\mu_2\mu_1'^2 + \mu_1'^4$$
$$= 142{\cdot}3125 + 4(39{\cdot}75)(2{\cdot}5) + 6(14{\cdot}75)(2{\cdot}5)^2 + (2{\cdot}5)^4$$
$$= 142{\cdot}3125 + 397{\cdot}5 + 553{\cdot}125 + 39{\cdot}0625$$
$$= 1132.$$

*Moments about the point $x = 2$.* We have $\bar{x} = A + \mu_1'$. Taking $A = 2$, the first moment about the point $x = 2$ is

$$\mu_1' = \bar{x} - 2 = 2{\cdot}5 - 2 = 0{\cdot}5$$

Hence

$$\mu_2' = \mu_2 + \mu_1'^2 = 14{\cdot}75 + 0{\cdot}25 = 15$$
$$\mu_3' = \mu_3 + 3\mu_2\mu_1' + \mu_1'^3 = 39{\cdot}75 + 3(14{\cdot}75)(0{\cdot}5) + (0{\cdot}5)^3$$
$$= 39{\cdot}75 + 22{\cdot}125 + 0{\cdot}125 = 62$$
$$\mu_4' = \mu_4 + 4\mu_3\mu_1' + 6\mu_2\mu_1'^2 + \mu_1'^4$$
$$= 142{\cdot}3125 + 4(39{\cdot}75)(0{\cdot}5) + 6(14{\cdot}75)(0{\cdot}5)^2 + (0{\cdot}5)^4$$
$$= 142{\cdot}3125 + 79{\cdot}5 + 22{\cdot}125 + 0{\cdot}0625$$
$$= 244$$

**Example 3·9.** *Calculate the first four moments of the following distribution about the mean and hence find* $\beta_1$ and $\beta_2$.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *x:* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| *f:* | 1 | 8 | 28 | 56 | 70 | 56 | 28 | 8 | 1 |

**Solution.** CALCULATION OF MOMENTS

| $x$ | $f$ | $d = x - 4$ | $fd$ | $fd^2$ | $fd^3$ | $fd^4$ |
|---|---|---|---|---|---|---|
| 0 | 1 | –4 | –4 | 16 | –64 | 256 |
| 1 | 8 | –3 | –24 | 72 | –216 | 648 |
| 2 | 28 | –2 | –56 | 112 | –224 | 448 |
| 3 | 56 | –1 | –56 | 56 | –56 | 56 |
| 4 | 70 | 0 | 0 | 0 | 0 | 0 |
| 5 | 56 | 1 | 56 | 56 | 56 | 56 |
| 6 | 28 | 2 | 56 | 112 | 224 | 448 |
| 7 | 8 | 3 | 24 | 72 | 216 | 648 |
| 8 | 1 | 4 | 4 | 16 | 64 | 256 |
| Total | 256 | 0 | 0 | 512 | 0 | 2,816 |

Moments about the points $x = 4$ are

$$\mu_1' = \frac{1}{N}\Sigma fd = 0, \quad \mu_2' = \frac{1}{N}\Sigma fd^2 = \frac{512}{256} = 2,$$

$$\mu_3' = \frac{1}{N}\Sigma fd^3 = 0 \text{ and } \mu_4' = \frac{1}{N}\Sigma fd^4 = \frac{2816}{256} = 11$$

Moments about mean are :

$$\mu_1 = 0, \quad \mu_2 = \mu_2' - \mu_1'^2 = 2$$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 = 0$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4 = 11$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0, \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{11}{4} = 2{\cdot}75$$

**Example 3·10** *For a distribution the mean is 10, variance is 16, $\gamma_2$ is + 1 and $\beta_2$ is 4. Obtain the first four moments about the orgin, i.e., zero. Comment upon the nature of distribution.*

**Solution.** We are given

Mean = 10, $\mu_2 = 16$, $\gamma_1 = +1$, $\beta_1 = 4$

*First four moments about origin* ($\mu_1'$, $\mu_2'$, $\mu_3'$, $\mu_4'$)

$\mu_1'$ = First moment about origin = Mean = 10

$$\mu_2 = \mu_2' - \mu_1'^2 \Rightarrow \mu_2' = \mu_2 + \mu_1'^2 \Rightarrow \mu_2' = 16 + 10^2 = 116$$

we have $\gamma_1 = +1 \Rightarrow \dfrac{\mu_3}{\mu_2^{3/2}} = 1$

$$\Rightarrow \mu_3 = \mu_2^{3/2} = (16)^{3/2} = 4^3 = 64$$

$$\therefore \mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$$

$$\Rightarrow \mu_3' = \mu_3 + 3\mu_2'\mu_1' - 2\mu_1'^3$$

$$= 64 + 3 \times 116 \times 10 - 2 \times 1000 = 3544 - 2000 = 1544$$

Now $\beta_2 = \dfrac{\mu_4}{\mu_2^2} = 4 \Rightarrow \mu_4 = 4 \times 16^2 = 1024$

and $\mu_4 = \mu'_4 - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - \mu_1'^4$

$\Rightarrow \quad \mu_4' = 1024 + 4 \times 1544 \times 10 - 6 \times 116 \times 100 + 3 \times 10000$
$= 92784 - 69600 = 23184.$

*Comments of Nature of distribution* : [c.f. § 3·13 and § 3·14] Since $\gamma_1 = +1$, the distribution is moderately positively skewed, i.e, if we draw the curve for the given distribution, it will have longer tail towards the right. Further since $\beta_2 = 4 > 3$, the distribution is leptokurtic, i.e., it will be more peaked than the normal curve.

**Example 3·11.** *If for a random variable x, the absolute moment of order k exists for ordinary $k = 1, 2, ..., n-1$, then the following inequalities*

(*i*) $\beta_k^{2k} \le \beta_{k-1}^k \cdot \beta_{k+1}^k$, (*ii*) $\beta_k^{1/k} \le \beta_{k+1}^{1/k+1}$

*holds for $k=1, 2, ..., n-1$, where $\beta_k$ is the kth absolute moment about the origin.*

**[Delhi Univ. B.Sc. (Stat.Hons.) 1989]**

**Solution.** If $x_i | f_i$, $i = 1, 2, ..., n$ is the given frequency distribution, then

$$\beta_k = \frac{1}{N} \Sigma f_i \left| x_i^k \right| \qquad ...(1)$$

Let $u$ and $v$ be arbitrary real numbers, then the expression

$$\sum_{i=1}^{n} f_i \left[ u \,| x_i^{(k-1)/2}| + v| x_i^{(k+1)/2}| \right]^2, \text{ is non-negative.}$$

$$\Rightarrow \sum_{i=1}^{n} f_i \left[ u \,| x_i |^{(k-1)/2} + v| x_i |^{(k+1)/2} \right]^2 \ge 0$$

$$\Rightarrow \mu^2 \Sigma f_i |x_i|^{k-1} + v^2 \Sigma f_i |x_i|^{k+1} + 2uv \Sigma f_i |x_i|^k \ge 0$$

Dividing throughout by N and using relation (1), we get

$$u^2 \beta_{k-1} + v^2 \beta_{k+1} + 2uv\beta_k \ge 0, \text{ i.e., } u^2 \beta_{k-1} + 2uv\beta_k + v^2 \beta_{k+1} \ge 0 \qquad ...(2)$$

We know that the condition for the expression $a x^2 + 2hxy + b y^2$ to be non - negative for all values of $x$ and $y$ is that

$$\begin{vmatrix} a & h \\ h & b \end{vmatrix} \ge 0$$

Using this result, we get from (2)

$$\begin{vmatrix} \beta_{k-1} & \beta_k \\ \beta_k & \beta_{k+1} \end{vmatrix} \ge 0$$

$$\Rightarrow \qquad \beta_{k-1} \cdot \beta_{k+1} - \beta_k^2 \ge 0 \qquad ...(3)$$

Raising both sides of (3) to power k, we get

$$\beta_k^{2k} \ge \beta_{k-1}^k \cdot \beta_{k+1}{}^k \qquad ...(4)$$

Putting k=1, 2, ..., k-1, k successively in (4), we get

$$\beta_1{}^2 \le \beta_0 \beta_2$$
$$\beta_2^4 \le \beta_1^2 \beta_3^2$$
$$\beta_3^6 \le \beta_2^3 \beta_4^3$$
$$\cdots\cdots$$
$$\cdots\cdots$$
$$\beta_{k-1}^{2(k-1)} \le \beta_{k-2}^{k-1} \cdot \beta_k^{k-1}$$

$$\beta_k^{2k} \leq \beta_{k-1}^k \beta_{k+1}^k$$

Multiplying these inequalities and noting that $\beta_0 = 1$, we get

$$\beta_k^{k+1} \leq \beta_{k+1}^k \text{ for } k = 1, 2, \ldots, n-1.$$

Raising both sides of the inequality to the power $\dfrac{1}{k(k+1)}$, we get

$$\beta_k^{1/k} \leq \beta_{k+1}{}^{1/(k+1)} \qquad \ldots (5)$$

**Remark.** Result (5) shows that $\beta_k{}^{1/k}$ is an increasing function of $k$.

## EXERCISE 3 (b)

**1.** (*a*) Define the raw and central moments of a frequency distribution. Obtain the relation between the central moments of order $r$ in terms of the raw moments. What are Sheppard's corrections to the central moments ?

(*b*) Define moments. Establish the relationship between the moments about mean, *i.e.,* Central Moments in terms of moments about any arbitrary point and *vice versa.*

The first three moments of a distribution about the value 2 of the variable are 1, 16 and $-$ 40. Show that the mean is 3, the variance is 15 and $\mu_3 = -86$. Also show that the first three moments about $x = 0$ are 3, 24 and 76.

(*c*) For a distribution the mean is 10, variance is 16, $\gamma_1$ is + 1 and $\beta_2$ is 4. Find the first four moments about the origin.

**Ans.** $\mu_1' = 10, \mu_2' = 116, \mu_3' = 1544$ and $\mu_4' = 23184$.

(*d*) (*i*) Define 'moment'. What is its use ? Express first four central moments in terms of moments about the origin. What is the effect of change of origin and scale on $\mu_3$ ?

(*ii*) The first three moments of a distribution about the point $X = 7$ are 3, 11 and 15 respectively. Obtain mean, variance and $\beta_1$.

**2.** The first four moments of distribution about the value 5 of the variable are 2, 20, 40 and 50. Obtain as far as possible, the various characteristics of the distribution on the basis of the information given.

**Ans.** Mean = 7, $\mu_2 = 16$, $\mu_3 = -64$, $\mu_4 = 162$, $\beta_1 = 1$ and $\beta_2 = 0·63$.

**3.** (*a*) If the first four moments of a distribution about the value 5 are equal to $-4$, 22, $-$ 117 and 560, determine the corresponding moments (*i*) about the mean, (*ii*) about zero.

(*b*) What is Sheppard's correction? What will be the corrections for the first four moments ?

The first four moments of a distribution about $x = 4$ are 1, 4, 10, 45. Show that the mean is 5 and the variance is 3 and $\mu_3$ and $\mu_4$ are 0 and 26 respectively,

(*c*) In certain distribution, the first four moments about the point 4 are $-1·5$, 17, $-$ 13 and 308. Calculate $\beta_1$ and $\beta_2$.

(*d*) The first four moments of a frequency distribution about the point 5 are $-0.55$, 4·46, $-$ 0·43 and 68·52. Find $\beta_1$ and $\beta_2$.

**Ans.** $\mu_2 = 4·1575$, $\mu_3 = 6·5962$, $\mu_4 = 75·3944$, $\beta_1 = 0·6055$, $\beta_2 = 4·3619$.

**4.** (*a*) For the following data, calculate (*i*) Mean, (*ii*) Median, (*iii*) Semi-inter-quartile range, (*iv*) Coefficient of variation, and (*v*) $\beta_1$ and $\beta_2$ coefficients.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Wages in Rupees : | 170—180 | 180—190 | 190—200 | 200—210 | 210—220 | 220—230 | 230—240 | 240—250 |
| No. of Persons : | 52 | 68 | 85 | 92 | 100 | 95 | 70 | 28 |

**Ans.** Mean = 209 (approx.); Median = 209·8; Q.D. = 15·8; $\sigma$ = 19·7; $C.V.$ = 9·4; $\beta_1$ = 0·003; $\beta_2$ = 26·105.

(*b*) Find the second, third and fourth central moments of the frequency distribution given below. Hence find (*i*) a measure of skewness ($\gamma_1$) and (*ii*) measure of kurtosis ($\gamma_2$).

| *Class Limits* | *Frequency* |
|---|---|
| 100·0 – 114·9 | 5 |
| 115·0 – 119·9 | 15 |
| 120·0 – 124·9 | 20 |
| 125·0 – 129·9 | 35 |
| 130·0 – 134·9 | 10 |
| 135·0 – 139·9 | 10 |
| 140·0 – 144·9 | 5 |

Also apply Sheppard's corrections for moments.

**Ans.** $\mu_2 = 2{\cdot}16,\ \mu_3 = 0{\cdot}804,\ \mu_4 = 12{\cdot}5232$

$$\gamma_1 = \sqrt{\beta_1} = 0{\cdot}25298;\ \gamma_2 = \beta_2 - 3 = -0{\cdot}317.$$

(*c*) The standard deviation of a symmetrical distribution is 5. What must be the value of the fourth moment about the mean in order that the distribution be (*i*) leptokurtic, (*ii*) mesokurtic, and (*iii*) platykurtic ?

**Hint :** $\mu_1 = \mu_3 = 0$ (because distribution is symmetrical).

$$\sigma = 5 \Rightarrow \sigma^2 = \mu_2 = 25$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{625}$$

(*i*) Distribution is leptokurtic if $\beta_2 > 3$, *i.e.*, if $\frac{\mu_4}{625} > 3 \Rightarrow \mu_4 > 1875$

(*ii*) Distribution is mesokurtic if $\beta_2 = 3 \Rightarrow$ if $\mu_4 = 1875$

(*iii*) Distribution is platykurtic if $\beta_2 < 3 \Rightarrow$ if $\mu_4 < 1875$

5. Show that for discrete distribution $\beta_2 > 1$.

**[Allahabad Univ. M.A., 1993; Delhi Univ. B.Sc. (Stat. Hons), 1992]**

**Hint.** We have to show that $\mu_4/\mu_2^2 > 1$, *i.e.*, $\mu_4 > \mu_2^2$. *If* $x_i\,/f_i$, $i = 1, 2, \ldots, n$, be the given discrete distribution, then we have to prove that

$$\frac{1}{N}\sum_i f_i\,(x_i - \bar{x})^4 > \left(\frac{1}{N}\sum_i f_i\,(x_i - \bar{x})^2\right)^2$$

Putting $(x_i - \bar{x})^2 = z_i$, we have to show that

$$\frac{1}{N}\sum_i f_i z_i^2 > \left(\frac{1}{N}\sum_i f_i z_i\right)^2$$

*i.e.,* $\frac{1}{N}\sum_i f_i z_i^2 - \left(\frac{1}{N}\sum_i f_i z_i\right)^2 > 0$

*i.e.,* $\sigma_z^2 > 0$,

which is always true, since variance is always positive.

Hence $\beta_2 > 1$.

**6.** (a) The scores in Economics of 250 candidates appearing at an examination have

| | |
|---|---|
| Mean | $= \bar{x} = 39{\cdot}72$ |
| Variance | $= \sigma^2 = 97{\cdot}80$ |
| Third Central moment | $= \mu_3 = -114{\cdot}18$ |
| Fourth central moment | $= \mu_4 = 28{,}396{\cdot}14$ |

It was later found on scrutiny that the score 61 of a candidate has been wrongly recorded as 51. Make necessary corrections in the given values of the mean and the central moment. **(Gujarat Univ. M.A., 1993)**

(b) For a distribution of 250 heights, calculations showed that the mean, standard deviation, $\beta_1$ and $\beta_2$ were 54 inches, 3 inches 0 and 3 inches respectively. It was, however, discovered on checking that the two items 64 and 50 in the original data were wrongly written in place of correct values 62 and 52 inches respectively. Calculate the correct frequency constants.

**Ans.** Correct Mean = 54, S.D. = 2·97, $\mu_3 = -2{\cdot}18$, $\mu_4 = 218{\cdot}42$, $\beta_1 = 0{\cdot}0070$ and $\beta_2 = 2{\cdot}81$

**7.** In calculating the moments of a frequency distribution based on 100 observations, the following results are obtained :

Mean = 9, Variance = 19, $\beta_1 = 0{\cdot}7$ ( $\mu_3$ + ive ), $\beta_2 = 4$

But later on its was found that one observation 12 was read as 21. Obtain the correct value of the first four central moments.

**Ans.** Corrected mean = 8·91, $\mu_2 = 17{\cdot}64$, $\mu_3 = 57{\cdot}05$, $\mu_4 = 1257{\cdot}15$, $\beta_1 = 0{\cdot}59$ and $\beta_2 = 4{\cdot}04$.

**8.** (a) Show that if a range of six times the standard deviation covers at least 18 class intervals, Sheppard's correction will make a difference of less then 0·5 percent in the corrected value of the standard deviation.

**Hint.** If $h$ is the magnitude of the class interval, then we want :

$$6\sigma > 18h \Rightarrow \sigma > 3h \Rightarrow h^2 < \frac{1}{9}\sigma^2 \Rightarrow -h^2 > -\frac{1}{9}\sigma^2$$

$$\therefore \quad \mu_2\,(\text{corrected}) = \mu_2 - \frac{h^2}{12} \geq \sigma^2 - \frac{1}{9 \times 12}\sigma^2 = \sigma^2\left(1 - \frac{1}{108}\right)$$

$$\Rightarrow \text{s.d. (corrected)} \geq \sigma\left(1 - \frac{1}{108}\right)^{1/2} \simeq \sigma\left(1 - \frac{1}{2} \times \frac{1}{108}\right)$$

$$\therefore \text{ Required adjustement} = \sigma - \sigma\,(\text{corrected}) < \frac{\sigma}{216} < \frac{\sigma}{200} = \frac{1}{2}\,\%\text{ of s.d.}$$

(*b*) Show that, if the class intervals of a grouped distribution is less than one-third of the calculated standard deviation, Sheppard's adjustment makes a difference of less than $\frac{1}{2}$ % in the estimate of the standard deviation

**9.** (*a*) If $\partial_r$ is the *r*th absolute moment about zero, use the mean value of

$$[u \mid x \mid^{(r-1)/2} + v \mid x \mid^{(r+1)/2}]^2$$

to show that

$$(\delta_r)^{2r} \le (\partial_{r-1})^r (\partial_{r+1})^r$$

From this derive the following inequalities :

(*i*) $(\partial_r)^{r+1} \le (\delta_{r+1})^r$, (*ii*) $(\partial_r)^{1/r} \le (\delta_{r+1})^{1/(r+1)}$

(*b*) For a random variable $X$ moments of all order exist. Denoting by $\mu_j$ and $\partial_j$, the *j*th central moment and *j*th absolute moment respectively, show that

(*i*) $(\mu_{2j+1})^2 \le \mu_{2j}\ \mu_{2j+2}$,

(*ii*) $(\partial_j)^{1/j} \le (\partial_{j+1})^{1/(j+1)}$ **(Karnataka Univ. B.Sc., 1993)**

**10.** If $\beta_1$ and $\beta_2$ are the Pearsons's coefficients of skewness and Kurtosis respectively, show that $\beta_2 > \beta_1 + 1$. **(Bangalore Univ. B.Sc., 1993)**

**3·13. Skewness.** Literally, skewness means '*lack of symmetry*'. We study skewness to have an idea about the shape of the curve which we can draw with the help of the given data. A distribution is said to be skewed if

(*i*) Mean, median and mode fall at different points,

*i.e.*, Mean ≠ Median ≠ Mode,

(*ii*) Quartiles are not equidistant from median, and

(*iii*) The curve drawn with the help of the given data is not symmetrical but stretched more to one side than to the other.

**Measures of Skewness.** Various measures of skewness are

(1) $S_k = M - M_d$ (2) $S_k = M - M_0$.

where $M$ is the mean, $M_d$, the median and $M_0$, the mode of the distribution.

(3) $S_k = (Q_3 - M_d) - (M_d - Q_1)$.

These are the absolute measures of skewness. As in dispersion, for comparing two series we do not calculate these absolute measures but we calculate the relative measures called the *co-efficients of skewness* which are pure numbers independent of units of measurement. The following are the *coefficients of Skewness.*

**1.** *Prof. Karl Pearson's Coefficient of Skewness.*

$$S_k = \frac{(M - M_0)}{\sigma} \qquad \ldots (3\cdot27)$$

where $\sigma$ is the standard deviation of the distribution.

If mode is ill-defined, then using the relation, $M_0 = 3M_d - 2M$, for a moderately asymmetrical distribution, we get

$$S_k = \frac{3(M - M_d)}{\sigma} \qquad \ldots (3\cdot27a)$$

The limits for Karl Pearson's coefficient of skewness are ± 3. In practice, these limits are rarely attained.

Skewness is positive if $M > M_0$ or $M > M_d$ and negative if $M < M_0$ or $M < M_d$.

II. *Prof. Bowley's Coefficient of Skewness.* Based on quartiles,

$$S_K = \frac{(Q_3 - M_d) - (M_d - Q_1)}{(Q_3 - M_d) + (M_d - Q_1)} = \frac{Q_3 + Q_1 - 2M_d}{Q_3 - Q_1} \qquad \ldots (3\cdot 28)$$

**Remarks 1.** Bowley's coefficient of skewness is also known as *Quartile coefficient of skewness* and is especially useful in situations where quartiles and median are used, *viz.*,

(*i*) When the mode is ill-defined and extreme observations are present in the data.

(*ii*) When the distribution has open end classes or unequal class intervals.

In these situations Pearson's coefficient of skewness cannot be used.

2. From (3·28), we observe that

$$S_k = 0 \text{ , if } Q_3 - M_d = M_d - Q_1$$

This implies that for a symmetrical distribution ($S_k = 0$), median is equidistant from the upper and lower quartiles. Moreover skewness is positive if :

$$Q_3 - M_d > M_d - Q_1 \;\Rightarrow\; Q_3 + Q_1 > 2M_d$$

and skewness is negative if

$$Q_3 - M_d < M_d - Q_1 \;\Rightarrow\; Q_3 + Q_1 < 2M_d$$

3. *Limits for Bowley's Coefficient of Skewness.* We know that for two real positive numbers $a$ and $b$ (*i.e.*, $a > 0$ and $b > 0$), the moduls value of the difference $(a - b)$ is always less than or equal to the modules value of the sum $(a + b)$, *i.e.*,

$$|a - b| \leq |a + b| \;\Rightarrow\; \left|\frac{a - b}{a + b}\right| \leq 1 \qquad \ldots (*)$$

We also know that $(Q_3 - M_d)$ and $(M_d - Q_1)$ are both non-negative. Thus, taking $a = Q_3 - M_d$ and $b = M_d - Q_1$, in (*), we get

$$\left|\frac{(Q_3 - M_d) - (M_d - Q_1)}{(Q_3 - M_d) + (M_d - Q_1)}\right| \leq 1$$

$$\Rightarrow \qquad |S_k \text{ (Bowley)}| \leq 1$$

$$\Rightarrow \qquad -1 \leq Sk \text{ (Bowley)} \leq 1.$$

Thus, Bowley's coefficient of skewness ranges from – 1 to 1.

Further, we note from (3·28) that :

$$S_k = +1, \text{ if } M_d - Q_1 = 0, \textit{ i.e., } \text{if } M_d = Q_1$$

$$S_k = -1, \text{ if } Q_3 - M_d = 0, \textit{ i.e., } \text{if } Q_3 = M_d.$$

4. It should be clearly understood that the values of the coefficients of skewness obtained by Bowley's formula and Pearson's formula are not comparable, although in each case, $S_k = 0$, implies the absence of skewness, *i.e.*, the distribution is symmetrical. It may even happen that one of them gives positive skewness while the other gives negative skewness.

5. In Bowley's coefficient of skewness the disturbing factor of variation is eliminated by dividing the absolute measure of skewness, viz., $(Q_3 - Md) - (Md - Q_1)$ by the measure of dispersion $(Q_3 - Q_1)$, i.e., quartile range.

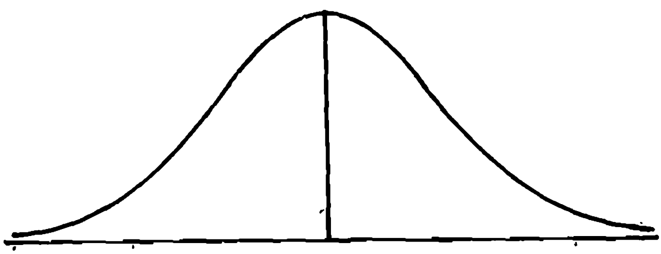
6. The only and perhaps quite serious limitations of this coefficient is that it is based only on the central 50% of the data and ignores the remaining 50% of the data towards the extremes.

**III. Based upon moments, co-efficient of skewness is**

$$S_k = \frac{\sqrt{\beta_1}\,(\beta_2 + 3)}{2\,(5\beta_2 - 6\beta_1 - 9)} \qquad \ldots(3{\cdot}29)$$

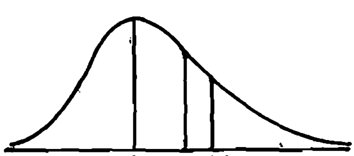
where symbols have their usual meaning. Thus $S_k = 0$ if either $\beta_1 = 0$ or $\beta_2 = -3$. But since $\beta_2 = \mu_4/\mu_2^2$, cannot be negative, $S_k = 0$ if and only if $\beta_1 = 0$. Thus for a symmetrical distribution $\beta_1 = 0$. In this respect $\beta_1$ is taken to be a measure of skewness. The co-efficient, in (3·29) is to be regarded as without sign.

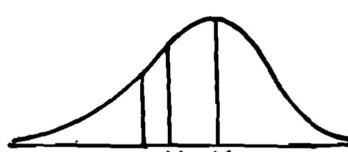
We observe in (3·27) and (3·28) that skewness can be positive as well as negative. The skewness is positive if the larger tail of the distribution lies towards



$\bar{x}$ (*Mean*) $= M_0 = M_d$
(Symmetrical Distribution)

the higher values of the variate (the right), i.e., if the curve drawn with the help of the given data is tretched more to the right than to the left and is negative



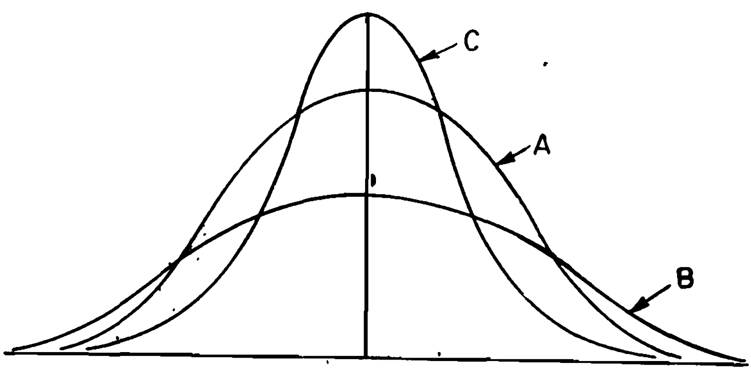




(Positively Skewed Distribution) (Negatively Skewed Distribution) in the contracy case.

**3·14. Kurtosis.** If we know the measures of central tendency, dispersion and skewness, we still cannot form a complete idea about the distribution as will be clear from the following figure in which all the three curves $A$, $B$ and $C$ are symmetrical about the mean '$m$' and have the same range.

In addition to these measures we should know one more measure which Prof. Karl Pearson calls as the '*Convexity of curve*' or *Kurtosis*. Kurtosis enables us to have an idea about the flatness or peakedness of the curve. It is measured by the co-efficient $\beta_2$ or its derivation $r_2$ given by

$$\beta_2 = \mu_4/\mu_2^2, \quad \gamma_2 = \beta_2 - 3$$




Curve of the type '$A$' which is neither flat nor peaked is called the *normal curve or mesokurtic curve* and for such a curve $\beta_2 = 3$, i.e., $\gamma_2 = 0$. Curve of the type '$B$' which is flatter than the normal curve is known as *platykurtic* and for such a curve $\beta_2 < 3$, *i.e.*, $\gamma_2 < 0$. Curve of the type '$C$' which is more peaked than the normal curve is called *leptokurtic* and for such a curve $\beta_2 > 3$, i.e., $\gamma_2 > 0$.

## EXERCISE 3 (c)

1. What do you understand by skewness ? How is it measured ? Distinguish clearly, by giving figures, between positive and negative skewness.

2. Explain the methods of measuring skewness and kurtosis of a frequency distribution.

3. Show that for any frequency distribution :

(i) Kurtosis is *greater* than unity.

(ii) Co-efficient of skewness is *less* than 1 numerically.

4. Why do we calculate in general, only the first four moments about mean of a distribution and not the higher moments ?

5. (a) Obtain Karl Pearsons's measure of skewness for the following data:

| *Values* | *Frequency* | *Values* | *Frequency* |
|---|---|---|---|
| 5 – 10 | 6 | 25 – 30 | 15 |
| 10 – 15 | 8 | 30 – 35 | 11 |
| 15 – 20 | 17 | 35 – 40 | 2 |
| 20 – 25 | 21 | | |

(b) Assume that a firm has selected a random sample of 100 from its production line and has obtain the data shown in the table below :

| *Class interval* | *Frequency* | *Class interval* | *Frequency* |
|---|---|---|---|
| 130 – 134 | 3 | 150 – 154 | 19 |
| 135 – 139 | 12 | 155 – 159 | 12 |
| 140 – 144 | 21 | 160 – 164 | 5 |
| 145 – 149 | 28 | | |

Compute the following :

(a) The arithmetic mean, (b) the standard deviation,

(c) Karl Pearson's coefficient of skewness.

**Ans.** (a) 147·2, (b) 7·2083 (c) 0·0711

**6.** (a) For the frequency distribution given below, calculate the coefficient of skewness based on quartiles.

| *Annual Sales* (*Rs.* '000 ) | *No. of Firms* | *Annual Sales* (*Rs.* '000 ) | *No. of firms* |
|---|---|---|---|
| Less than 20 | 30 | Less than 70 | 644 |
| Less than 30 | 225 | Less than 80 | 650 |
| Less than 40 | 465 | Less than 90 | 665 |
| Less than 50 | 580 | Less than 100 | 680 |
| Less than 60 | 634 | | |

(b) (i) Karl Pearsons's coefficient of skewness of a distribution is 0·32, its s.d. is 6·5 and mean is 29·6. Find the mode of the distribution.

(ii) If the mode of the above distribution is 24·8, what will be the s.d. ?

**7.** (a) In a frequency distribution, the co-efficient of skewness based upon the quartiles is 0·6. If the sum oif the upper and lower quartiles is 100 and median is 38, find the value of the upper and lower quartiles.

**Hint.** We are given

$$S_k = \frac{Q_3 + Q_1 - 2\ Md}{Q_3 - Q_1} = 0{\cdot}6 \qquad \ldots(*)$$

Also $Q_3 + Q_1 = 100$ and Median = 38

Subsituting (*), we get

$$\frac{100 - 2 \times 38}{Q_3 - Q_1} = 0{\cdot}6$$

$$\Rightarrow \qquad Q_3 - Q_1 = 40$$

Simplifying we get $\quad Q_1 = 30,\ Q_3 = 70$

(b) A frequency distribution gives the following results :

(i) C.V. = 5 (ii) Karl Pearsons's co-efficient of skewness = 0·5

(iii) $\sigma = 2$.

Find the mean and mode of the distribution.

(c) find the C.V. of a frequency distribution given that its mean is 120, mode is 123 and Karl Pearons's co-efficient of skewness is – 0·3.

Ans. C.V. = 8·33

(d) The first three moments of distribution about the value 2 are 1, 16 and 40 respectively. Examine the skewness of the distribution.

8. The first three moments about the origin 51 Kg. calculated from the data on the weights of 25 college students are

$$\mu_1' = +\ 0{\cdot}4 \text{ kg.}, \sqrt{\mu_2'} = 1{\cdot}2 \text{ kg. and } (\mu_3')^{1/2} = -\ 0{\cdot}25\text{kg.}$$

Determine the mean, the standard deviation and coefficient oif skewness.

9. The first three moments about the origin are given by

$$\mu_1' = \frac{n+1}{2},\ \mu_2' = \frac{(n+1)(2n+1)}{6} \text{ and } \mu_3' = \frac{n\ 9n+1)^2}{4}$$

Examine the skewness of the data.

10. Find out the kurtosis of the data given below :

| *Class interval* | 0 – 10 | 10 – 20 | 20 – 30 | 30 – 40 |
|---|---|---|---|---|
| *Frequency* | 1 | 3 | 4 | 2. |

11. Data were obtained for distribution of passengers, entering Bombay local trains over time at intervals of 15 minutes for morning and evening rush hours separately, and the following results were obtained.

| | *Morning hours* | *Evening hours* |
|---|---|---|
| Arithmetic mean (Peak Hours) | 8 hrs. 38 min. | 5 hrs. 40 min. |
| Standard deviation | 38·5 min. | 34·9 min. |
| Coefficient of skewness (in 15 min. unit) | - 0·32 | + 0·17 |
| Kurtosis measure | 2·0 | 2·2 |

Interpret the result and discuss giving reasons, whether you approve of the measure of 'peak hour'.

12. (a) The standard deviation of a symmertrical distribution is 5., What must be the value of the fourth moment about the mean in order that the distribution be (i) Leptokurtic, (ii) mesokurtic, and (iii) platykurtic.

**Hint.** $\mu_1 = \mu_3 = 0$ (Because distribution is symmetrical), $\sigma = 5 \Rightarrow \sigma^2 = \mu_2 = 25$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{625}$$

(i) Distt. is leptokurtic if $\beta_2 > 3$ i.e., if $\frac{\mu_4}{625} > 3 \Rightarrow \mu_4 > 1875$

(ii) Distt. is mesokurtic if $\beta_2 = 3 \Rightarrow$ if $\mu_4 < 1875$

(*iii*) Distt. is platykurtic if $\beta_2 < 3 \Rightarrow$ if $\mu_4 < 1875$.

(*b*) Find the second, third and fourth central moments of the frequency distribution given below., Hence, find (*i*) a measure of skewness, and (*ii*) a measure of kurtosis ($\gamma_2$).

| *Class limits* | *Frequency* |
|---|---|
| 110·0 — 114·9 | 5 |
| 115·0— 119·9 | 15 |
| 120·0— 124·9 | 20 |
| 125·0— 129·9 | 35 |
| 130·0— 134·9 | 10 |
| 135·0— 139·9 | 10 |
| 140·0— 144·9 | 5 |

**Ans.** $\mu_2 = 2{\cdot}16, \mu_3 = 0{\cdot}804, \mu_4 = 12{\cdot}5232.$

$$\gamma_1 = \sqrt{\beta_1} = 0{\cdot}25298 \; ; \gamma_2 = \beta_2 - 3 = -\ 0{\cdot}317.$$

**13.** (*a*) Define Pearsonian coefficients $\beta_1$ and $\beta_2$ and discuss their utility in statistics. **[Delhi Univ. B.Sc. (Hons.), 1993]**

(*b*) What do you mean by skewness and kurtosis of a distribution ? Show that the Pearson's Beta coefficients satisfy the inequality $\beta_2 - \beta_1 - 1 \geq 0$. Also deduce that $\beta_2 \geq 1$. **(Delhi Univ. B.Sc. (Stat. Hons.), 1991)**

(*c*) Define the Pearson's coefficients $\gamma_1$ and $\gamma_2$ and discuss their utility in Statistics.

## OBJECTIVE TYPE QUESTIONS

**1.** Match the correct parts to make a valid statement.

| | |
|---|---|
| (*a*) Range | (*i*) $(Q_3 - Q_1)/2$ |
| (*b*) Quartile Deviation | (*ii*) $\sqrt{\frac{1}{N} \Sigma f_i (x_i - \bar{x})^2}$ |
| (*c*) Mean Deviation | (*iii*) $\frac{S.D.}{Mean} \times 100$ |
| (*d*) Standard Deviation | (*iv*) $\frac{1}{N} \Sigma f_i \mid (x_i - \bar{x}) \mid$ |
| (*e*) Coefficient of Variation | (*v*) $X_{max} - X_{min}$ |

**II.** Which value of '*a*' gives the minimum ?

(*i*) Mean square deviation from '*a*'

(*ii*) Mean deviation from '*a*'

**III.** Mean of 100 observations is 50 and S.D. is 10. What will be the new mean and S.D., if

(*i*) 5 is added to each observation,

(*ii*) each observation is multiplied by 3,

(*iii*) 5 is subtracted from each observation and then it is divided by 4?

**IV.** Fill in the blanks :

(*i*) (*a*) Absolute sum of deviation is minimum from...........

(b) Least value of root mean square deviation is .................

(ii) The sum of squares of deviations is least when measured from .................

(iii) The sum of 10 items is 12 and the sum of their squares is 16·9.

(iv) In any distribution, the standard deviation is always ................. the mean deviation.

(v) The relationship between root mean square deviation and standard deviation $\sigma$ is .................

(vi) If 25% of the items are less than 10 and 25% are more than 40, the coefficient of quartile deviation is .................

(vii) The median and standard deviation of a distribution are 20 and 4 respectively. If each item is increased by 2, the median will be ................. and the new standard deviation will be .................

(viii) In a symmetric distribution, the mean and the mode are .................

(xi) In symmetric distribution, the upper and the lower quartiles are equi-distant from .................

(x) If the mean, mode and standard deviation of a frequency distribution are 41, 45 and 8 respectively, then its pearson's coefficient of skewness is .................

(xi) For a symmetrical distribution $\beta_1$ = .................

(xii) If $\beta_2 > 3$ the distribution is said to be .................

(xiii) For a symmetric distribution $\mu_2$ = .................

$\mu_{2n+1}$ = .................

(xiv) If the mean and the mode of a given distribution are equal then its coefficient of skewness is .................

(xv) If the kurtosis of a distribution is 3, it is called ................. distribution.

(xvi) In a perfectly symmetrical distribution 50% of items are above 60 and 75% items are below 75. Therefore, the coefficient of quartile deviation is ................. and coefficient of skewness is .................

(xvii) Relation between $\beta_1$ and $\beta_2$ is given by .................

V. For the following questions given correct answers :

(i) Sum of absolute deviations about median is

(a) Least, (b) greatest, (c) zero, (d) equal.

(ii) The sum of squares of deviations is least when measured from

(a) Median, (b), (c) Mean, (d) Mode, (e) none of them.

(iii) In any discrete series (when all the values are not same) the relationship between M.D. about mean and S.D. is

(a) M.D. = S.D., (b) M.D. $\geq$ S.D., (c) M.D. < S.D.,

(d) M.D. $\leq$ S.D.

(e) None of these.

(iv) If each of a set of observations of a variable is multiplied by a constant (non-zero) value, the variance of the resultant variable.

(a) is unaltered, (b) increases (c) decreases, (d) is unknown.

(v) The appropriate measure whenever the extreme items are to be disregarded and when the distribution contains indefinite classes at the end is

(a) Median, (b) Mode, (c) Quartile deviation,
(d) Standard Deviation

(vi) A.M., G.M. and H.M. in any series are equal when

(a) the distribution is symmetric, (b) all the values are same,
(c) the distribution is positively skewed,
(d) the distribution is unimodal.

(vii) The limits for quartile coefficient of skewness are

(a) $\pm$ 3, (b) 0 and 3, (c) $\pm$ 1, (d) $\pm \infty$

(viii) 'The statement that the variance is equal to the second central moment'

(a) always true, (b) sometimes true, (c) never true,
(d) ambiguous.

(ix) The standard deviation of a distribution is 5. The value of the fourth central moment ( $\mu_4$ ), in order that the distribution be mesokurtic should be

(a) equal to 3, (b) greater than 1,875, (c) equal to 1,875,
(d) less than 1,875.

(x) In a frequency curve of scores the mode was found to be higher than the mean. This shows that the distribution is

(a) Symmetric, (b) negatively skewed, (c) positvely skewed,
(d) normal.

(xi) For any frequency distribution, the kurtosis is

(a) greater than 1, (b) less than 1, (c) equal to 1.

(xii) The measure of kurtosis is

(a) $\beta_2 = 0$, (b) $\beta_2 = 3$, (c) $\beta_2 = 4$.

(xiii) For the distribution

(a) $\mu_4 = 0$, (b) Median $= 0$,
(c) The distribution of $x$ is symmetrical.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $X$: | $-4$ | $-3$ | $-2$ | $-1$ | 0 | 1 | 2 | 3 | 4 | Total |
| $f$: | 2 | 4 | 5 | 7 | 10 | 7 | 5 | 4 | 2 | 46 |

(xiv) For a symmetric distribution

(a) $\mu_2 = 0$, (b) $\mu_2 > 0$; (c) $\mu_3 > 0$

**VI.** State which of the following statements are Ture and which False. In each of false statements given the correct statement.

(i) Mean, standard deviation and varaince have the same unit.

(ii) Standard deviation of every distribution is unique and always exists.

(iii) Median is the value of the variance which divides the total frequency it two equal parts.

(iv) Mean – Mode = 3 (mean – median) is often approximately satisified.

(v) Mean deviation = $\frac{4}{5}$ (standard deviation) is always satisfied.

(vi) $\beta_2 \geq 1$ is always satisfied

(vii) $\beta_1 = 0$ is a conclusive test for a distribution to be symmetrical.

CHAPTER FOUR

# *Theory of Probability*

**4·1. Introduction.** If an experiment is repeated under essentially homogeneous and similar conditions we generally come across two types of situations:

*(i)* The result or what is usually known as the '*outcome*' is unique or certain.

*(ii)* The result is not unique but may be one of the several possible outcomes.

The phenomena covered by *(i)* are known as '*deterministic*' or '*predictable*' phenomena. By a deterministic phenomenon we mean one in which the result can be predicted with certainty. For example :

*(a)* For a perfect gas,

$$V \propto \frac{1}{P} \quad i.e., \; PV = constant,$$

provided the temperature remains the same.

*(b)* The velocity '$v$' of a particle after time '$t$' is given by

$$v = u + at$$

where $u$ is the initial velocity and $a$ is the acceleration. This equation uniquely determines $v$ if the right-hand quantities are known.

*(c)* Ohm's Law, *viz.,* $C = \dfrac{E}{R}$

where C is the flow of current, E the potential difference between the two ends of the conductor and R the resistance, uniquely determines the value C as soon as E and R are given.

A deterministic model is defined as a model which stipulates that the conditions under which an experiment is performed determine the outcome of the experiment. For a number of situations the deterministic model suffices. However, there are phenomena [as covered by *(ii)* above] which do not lend themselves to deterministic approach and are known as '*unpredictable*' or '*probabilistic*' phenomena. For example :

*(i)* In tossing of a coin one is not sure if a head or tail will be obtained.

*(ii)* If a light tube has lasted for $t$ hours, nothing can be said about its further life. It may fail to function any moment.

In such cases we talk of chance or probability which is taken to be a quantitative measure of certainty.

**4·2. Short History.** Galileo (1564-1642), an Italian mathematician, was the first to attempt at a quantitative measure of probability while dealing with some problems related to the theory of dice in gambling. But the first foundation of the mathematical theory if probability was laid in the mid-seventeenth century by two French mathematicians, B. Pascal (1623-1662) and P. Fermat (1601-1665), while solving a number of problems posed by French gambler and noble man Chevalier-De-Mere to Pascal. The famous *'problem of points'* posed by De-Mere to Pascal is : "Two persons play a game of chance. The person who first gains a certain number of points wins the stake. They stop playing before the game is completed. How is the stake to be decided on the basis of the number of points each has won?" The two mathematicians after a lengthy correspondence between themselves ultimately solved this problem and this correspondence laid the first foundation of the science of probability. Next stalwart in this field was J. Bernoulli (1654-1705) whose 'Treatise on Probability' was published posthumously by his nephew N. Bernoulli in 1713. De-Moivre (1667-1754) also did considerable work in this field and published his famous 'Doctrine of Chances' in 1718. Other main contributors are : T. Bayes (Inverse probability), P.S. Laplace (1749-1827) who after extensive research over a number of years finally published 'Theoric analytique des probabilities' in 1812. In addition to these, other outstanding contributors are Levy, Mises and R.A. Fisher.

Russian mathematicians also have made very valuable contributions to the modern theory of probability. Chief contributors, to mention only a few of them are : Chebyshev (1821-94) who founded the Russian School of Statisticians; A. Markoff (1856-1922); Liapounoff (Central Limit Theorem); A. Khintchine (Law of Large Numbers) and A. Kolmogorov, who axiomised the calculus of probability.

**4·3. Definitions of Various Terms.** In this section we will define and explain the various terms which are used in the definition of probability.

**Trial and Event.** Consider an experiment which, though repeated under essentially identical conditions, does not give unique results but may result in any one of the several possible outcomes. The experiment is known as a *trial* and the outcomes are known as *events* or *cases*. For example :

*(i)* Throwing of a die is a trial and getting 1(or 2 or 3, ...or 6) is an event.

*(ii)* Tossing of a coin is a trial and getting head ($H$) or tail ($T$) is an event.

*(iii)* Drawing two cards from a pack of well-shuffled cards is a trial and getting a king and a queen are events.

**Exhaustive Events.** The total number of possible outcomes in any trial is known as exhaustive events or exhaustive cases. For example :

*(i)* In tossing of a coin there are two exhaustive cases, *viz.*, head and tail (the possibility of the coin standing on an edge being ignored).

*(ii)* In throwing of a die, there are six exhaustive cases since any one of the 6 faces 1, 2, ...,6 may come uppermost.

*(iii)* In drawing two cards from a pack of cards the exhaustive number of cases is $^{52}C_2$, since 2 cards can be drawn out of 52 cards in $^{52}C_2$ ways.

*(iv)* In throwing of two dice, the exhaustive number of cases is $6^2 = 36$, since any of the 6 numbers 1 to 6 on the first die can be associated with any of the six numbers on the other die.

In general in throwing of $n$ dice the exhaustive number of cases is $6^n$.

**Favourable Events or Cases.** The number of cases favourable to an event in a trial is the number of outcomes which entail the happening of the event. For example,

(i) In drawing a card from a pack of cards the number of cases favourable to drawing of an ace is 4, for drawing a spade is 13 and for drawing a red card is 26.

(ii) In throwing of two dice, the number of cases favourable to getting the sum 5 is : (1,4) (4,1) (2,3) (3,2), *i.e.*, 4.

**Mutually exclusive events.** Events are said to be *mutually exclusive* or *incompatible* if the happening of any one of them precludes the happening of all the others, *i.e.*, if no two or more of them can happen simultaneously in the same trial. For example :

(i) In throwing a die all the 6 faces numbered 1 to 6 are mutually exclusive since if any one of these faces comes, the possibility of others, in the same trial, is ruled out.

(ii) Similarly in tossing a coin the events head and tail are mutually exclusive.

**Equally likely events.** Outcomes of a trial are set to be equally likely if taking into consideration all the relevant evidences, there is no reason to expect one in preference to the others. For example

(i) In tossing an unbiased or uniform coin, head or tail are equally likely events.

(ii) In throwing an unbiased die, all the six faces are equally likely to come.

**Independent events.** Several events are said to be independent if the happening (or non-happening) of an event is not affected by the supplementary knowledge concerning the occurrence of any number of the remaining events. For example

(i) In tossing an unbiased coin the event of getting a head in the first toss is independent of getting a head in the second, third and subsequent throws.

(ii) If we draw a card from a pack of well-shuffled cards and replace it before drawing the second card, the result of the second draw is independent of the first draw. But, however, if the first card drawn is not replaced then the second draw is dependent on the first draw.

### 4·3·1. Mathematical or Classical or 'a priori' Probabality

*Definition.* If a trial results in $n$ exhaustive, mutually exclusive and equally likely cases and $m$ of them are favourable to the happening of an event $E$, then the probability '$p$' of happening of $E$ is given by

$$p = P(E) = \frac{\textit{Favourable number of cases}}{\textit{Exhaustive number of cases}} = \frac{m}{n} \qquad \text{....(4·1)}$$

Sometimes we express (4·1) by saying that 'the odds in favour of $E$ are $m : (n-m)$ or the odds against $E$ are $(n-m) : n$.'

Since the number of cases favourable to the 'non-happening' of the event $E$ are $(n-m)$, the probability '$q$' that $E$ will not happen is given by

$$q=\frac{n-m}{n}=1-\frac{m}{n}=1-p \quad \Rightarrow \quad p+q=1 \qquad \text{....(4·1}a)$$

Obviously $p$ as well as $q$ are non-negative and cannot exceed unity, *i.e*, $0 \le p \le 1$, $0 \le q \le 1$.

**Remarks. 1.** Probability '$p$' of the happening of an event is also known as the probability of success and the probability '$q$' of the non- happening of the event as the probability of failure.

2. If $P(E) = 1$, $E$ is called a *certain event* and if $P(E) = 0$, $E$ is called an *impossible* event.

3. **Limitations of Classical Definition.** This definition of Classical Probability breaks down in the following cases :

*(i)* If the various outcomes of the trial are not equally likely or equally probable. For example, the probability that a candidate will pass in a certain test is not 50% since the two possible outcomes, *viz.*, sucess and failure (excluding the possibility of a compartment) are not equally likely.

*(ii)* If the exhaustive number of cases in a trial is infinite.

### 4·3·2. Statistical or Empirical Probability

***Definition* (Von Mises).** *If a trial is repeated a number of times under essentially homogeneous and identical conditions, then the limiting value of the ratio of the number of times the event happens to the number of trials, as the number of trials become indefinitely large, is called the probability of happening of the event. (It is assumed that the limit is finite and unique).*

Symbolically, if in $n$ trials an event $E$ happens $m$ times, then the probability '$p$' of the happening of $E$ is given by

$$p = P(E) = \lim_{n \to \infty} \frac{m}{n} \qquad \text{....(4·2)}$$

**Example 4·1.** *What is the chance that a leap year selected at random will contain 53 Sundays?*

**Solution.** In a leap year (which consists of 366 days) there are 52 complete weeks and 2 days over. The following are the possible combinations for these two 'over' days:

*(i)* Sunday and Monday, *(ii)* Monday and Tuesday, *(iii)* Tuesday and Wednesday, *(iv)* Wednesday and Thursday, *(v)* Thursday and Friday, *(vi)* Friday and Saturday, and *(vii)* Saturday and Sunday.

In order that a leap year selected at random should contain 53 Sundays, one of the two 'over' days must be Sunday. Since out of the above 7 possibilities, 2 *viz.*, *(i)* and *(vii)*, are favourable to this event,

$$\therefore \qquad \text{Required probability} = \frac{2}{7}$$

**Example 4·2.** *A bag contains 3 red, 6 white and 7 blue balls. What is the probability that two balls drawn are white and blue?*

**Solution.** Total number of balls = 3 + 6 + 7 = 16.

Now, out of 16 balls, 2 can be drawn in ${}^{16}C_2$ ways.

$\therefore$ Exhaustive number of cases $= {}^{16}C_2 = \dfrac{16 \times 15}{2} = 120.$

Out of 6 white balls 1 ball can be drawn in ${}^{6}C_1$ ways and out of 7 blue balls 1 ball can be drawn in ${}^{7}C_1$ ways. Since each of the former cases can be associated with each of the latter cases, total number of favourable cases is : ${}^{6}C_1 \times {}^{7}C_1 = 6 \times 7 = 42.$

$$\therefore \qquad \text{Required probability} = \frac{42}{120} = \frac{7}{20}.$$

**Example 4·3.** (a) *Two cards are drawn at random from a well-shuffled pack of 52 cards. Show that the chance of drawing two aces is 1/221.*

*(b) From a pack of 52 cards, three are drawn at random. Find the chance that they are a king, a queen and a knave.*

*(c) Four cards are drawn from a pack of cards. Find the probability that (i) all are diamond, (ii) there is one card of each suit, and (iii) there are two spades and two hearts.*

**Solution.** *(a)* From a pack of 52 cards 2 cards can be drawn in ${}^{52}C_2$ ways, all being equally likely.

$\therefore$ Exhaustive number of cases $= {}^{52}C_2$

In a pack there are 4 aces and therefore 2 aces can be drawn in ${}^{4}C_2$ ways.

$$\therefore \qquad \text{Required probability} = \frac{{}^{4}C_2}{{}^{52}C_2} = \frac{4 \times 3}{2} \times \frac{2}{52 \times 51} = \frac{1}{221}$$

*(b)* Exhaustive number of cases $= {}^{52}C_3$

A pack of cards contains 4 kings, 4 queens and 4 knaves. A king, a queen and a knave can each be drawn in ${}^{4}C_1$ ways and since each way of drawing a king can be associated with each of the ways of drawing a queen and a knave, the total number of favourable cases $= {}^{4}C_1 \times {}^{4}C_1 \times {}^{4}C_1$

$$\therefore \qquad \text{Required probability} = \frac{{}^{4}C_1 \times {}^{4}C_1 \times {}^{4}C_1}{{}^{52}C_3} = \frac{4 \times 4 \times 4 \times 6}{52 \times 51 \times 50} = \frac{16}{5525}$$

*(c)* Exhaustive number of cases $= {}^{52}C_4$.

*(i)* $\qquad$ Required probability $= \dfrac{{}^{13}C_4}{{}^{52}C_4}$

*(ii)* $\qquad$ Required probability $= \dfrac{{}^{13}C_1 \times {}^{13}C_1 \times {}^{13}C_1 \times {}^{13}C_1}{{}^{52}C_4}$

*(iii)* $\qquad$ Required probability $= \dfrac{{}^{13}C_2 \times {}^{13}C_2}{{}^{52}C_4}$

**Example 4·4.** *What is the probability of getting 9 cards of the same suit in one hand at a game of bridge?*

**Solution.** One hand in a game of bridge consists of 13 cards.

$\therefore$ Exhaustive number of cases $= {}^{52}C_{13}$

Number of ways in which, in one hand, a particular player gets 9 cards of one suit are ${}^{13}C_9$ and the number of ways in which the remaining 4 cards are of some other suit are ${}^{39}C_4$. Since there are 4 suits in a pack of cards, total number of favourable cases $= 4 \times {}^{13}C_9 \times {}^{39}C_4$.

$$\therefore \qquad \text{Required probability} = \frac{4 \times {}^{13}C_9 \times {}^{39}C_4}{{}^{52}C_{13}}$$

**Example 4·5.** *(a) Among the digits 1, 2, 3, 4, 5, at first one is chosen and then a second selection is made among the remaining four digits. Assuming that all twenty possible outcomes have equal probabilities, find the probability that an odd digit will be selected (i) the first time, (ii) the second time, and (iii) both times.*

*(b) From 25 tickets, marked with the first 25 numerals, one is drawn at random. Find the chance that*

*(i) it is a multiple of 5 or 7,*

*(ii) it is a multiple of 3 or 7.*

**Solution.** *(a)* Total number of cases $= 5 \times 4 = 20$

*(i)* Now there are 12 cases in which the first digit drawn is odd, *viz.*, (1, 2), (1, 3), (1, 4), (1, 5), (3, 1), (3, 2), (3, 4), (3, 5), (5, 1), (5, 2), (5, 3) and (5, 4).

$\therefore$ The probability that the first digit drawn is odd

$$= \frac{12}{20} = \frac{3}{5}$$

*(ii)* Also there are 12 cases in which the second digit drawn is odd, *viz.*, (2, 1), (3, 1), (4, 1), (5, 1), (1, 3), (2, 3), (4, 3), (5, 3), (1, 5), (2, 5), (3, 5) and (4, 5).

$\therefore$ The probability that the second digit drawn is odd

$$= \frac{12}{20} = \frac{3}{5}$$

*(iii)* There are six cases in which both the digits drawn are odd, *viz.*, (1, 3), (1, 5), (3, 1), (3, 5), (5, 1) and (5, 3).

$\therefore$ The probability that both the digits drawn are odd

$$= \frac{6}{20} = \frac{3}{10}$$

*(b) (i)* Numbers (out of the first 25 numerals) which are multiples of 5 are 5, 10, 15, 20 and 25, *i.e.*, 5 in all and the numbers which are multiples of 7 are 7, 14 and 21, *i.e.*, 3 in all. Hence required number of favourable cases are 5+3=8.

$\therefore$ Required probability $= \dfrac{8}{25}$

*(ii)* Numbers (among the first 25 numerals) which are multiples of 3 are 3, 6, 9, 12, 15, 18, 21, 24, *i.e.*, 8 in all, and the numbers which are multiples of 7 are 7,

14, 21, *i.e.*, 3 in all. Since the number 21 is common in both the cases, the required number of distinct favourable cases is 8 + 3 − 1 = 10.

$$\therefore \qquad \text{Required probability} = \frac{10}{25} = \frac{2}{5}$$

**Example 4·6.** *A committee of 4 people is to be appointed from 3 officers of the production department, 4 officers of the purchase department, two officers of the sales department and 1 chartered accountant. Find the probability of forming the committee in the following manner:*

*(i) There must be one from each category.*

*(ii) It should have at least one from the purchase department.*

*(iii) The chartered accountant must be in the committee.*

**Solution.** There are 3+4+2+1=10 persons in all and a committee of 4 people can be formed out of them in ${}^{10}C_4$ ways. Hence exhaustive number of cases is

$${}^{10}C_4 = \frac{10 \times 9 \times 8 \times 7}{4\,!} = 210$$

*(i)* Favourable number of cases for the committee to consist of 4 members, one from each category is :

$${}^4C_1 \times {}^3C_1 \times {}^2C_1 \times 1 = 4 \times 3 \times 2 = 24$$

$$\therefore \qquad \text{Required probability} = \frac{24}{210} = \frac{8}{70}$$

*(ii)* $P$ [Committee has at least one purchase officer]

$$= 1 - P \text{ [Committee has no purchase officer]}$$

In order that the committee has no purchase officer, all the 4 members are to be selected from amongst officers of production department, sales department and chartered accountant, *i.e.*, out of 3+2+1=6 members and this can be done in ${}^5C_4 = \dfrac{6 \times 5}{1 \times 2} = 15$ ways. Hence

$$P \text{ [ Committee has no purchase officer ]} = \frac{15}{210} = \frac{1}{14}$$

$$\therefore \quad P \text{ [ Committee has at least one purchase officer ]} = 1 - \frac{1}{14} = \frac{13}{14}$$

*(iii)* Favourable number of cases that the committee consists of a chartered accountant as a member and three others are :

$$1 \times {}^9C_3 = \frac{9 \times 8 \times 7}{1 \times 2 \times 3} = 84 \text{ ways,}$$

since a chartered accountant can be selected out of one chartered accountant in only 1 way and the remaining 3 members can be selected out of the remaining 10 − 1 = 9 persons in ${}^9C_3$ ways. Hence the required probability $= \dfrac{84}{210} = \dfrac{2}{5}$.

**Example 4·7.** *(a) If the letters of the word 'REGULATIONS' be arranged at random, what is the chance that there will be exactly 4 letters between R and E?*

*(b) What is the probability that four S's come consecutively in the word 'MISSISSIPPI'?*

**Solution.** *(a)* The word '*REGULATIONS*' consists of 11 letters. The two letters $R$ and $E$ can occupy ${}^{11}P_2$, *i.e.*, $11 \times 10 = 110$ positions.

The number of ways in which there will be exactly 4 letters between $R$ and $E$ are enumerated below:

*(i)* $R$ is in the 1st place and $E$ is in the 6th place.

*(ii)* $R$ is in the 2nd place and $E$ is in the 7th place.

... ... ...

... ... ...

... ... ...

*(vi)* $R$ is in the 6th place and $E$ is in the 11th place.

Since $R$ and $E$ can interchange their positions, the required number of favourable cases is $2 \times 6 = 12$

$\therefore$ The required probability $= \dfrac{12}{110} = \dfrac{6}{55}$.

*(b)* Total number of permutations of the 11 letters of the word '*MISSISSIPPI*', in which 4 are of one kind *(viz., S)*, 4 of other kind *(viz., I)*, 2 of third kind *(viz., P)* and 1 of fourth kind *(viz., M)* are

$$\frac{11\,!}{4\,!\;4\,!\;2\,!\;1\,!}$$

Following are the 8 possible combinations of 4 $S$'s coming consecutively:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *(i)* | $S$ | $S$ | $S$ | $S$ | | | | | | |
| *(ii)* | — | $S$ | $S$ | $S$ | $S$ | | | | | |
| *(iii)* | — | — | $S$ | $S$ | $S$ | $S$ | | | | |
| ⋮ | | | ⋮ | | | ⋮ | | | | |
| *(viii)* | — | — | — | — | — | — | — | $S$ | $S$ | $S$ | $S$ |

Since in each of the above cases, the total number of arrangements of the remaining 7 letters, *viz.*, *MIIIPPI* of which 4 are of one kind, 2 of other kind and one of third kind are $\dfrac{7\,!}{4\,!\;2\,!\;1\,!}$, the required number of favourable cases

$$= \frac{8 \times 7\,!}{4\,!\;2\,!\;1\,!}$$

$$\therefore \quad \text{Required probability} = \frac{8 \times 7\,!}{4\,!\;2\,!\;1\,!} \div \frac{11\,!}{4\,!\;4\,!\;2\,!\;1\,!}$$

$$= \frac{8 \times 7\,! \times 4\,!}{11\,!} = \frac{4}{165}$$

**Example 4·8.** *Each coefficient in the equation $ax^2 + bx + c = 0$ is determined by throwing an ordinary die. Find the probability that the equation will have real roots.* **[Madras Univ. B. Sc. (Stat. Main), 1992]**

**Solution.** The roots of the equation $ax^2 + bx + c = 0$ ...(*)
will be real if its discriminant is non-negative, *i.e.*, if

$$b^2 - 4ac \geq 0 \quad \Rightarrow \quad b^2 \geq 4ac$$

Since each co-efficient in equation (*) is determined by throwing an ordinary die, each of the co-efficients $a$, $b$ and $c$ can take the values from 1 to 6.

$\therefore$ Total number of possible outcomes (all being equally likely)

$$= 6 \times 6 \times 6 = 216$$

The number of favourable cases can be enumerated as follows:

| *ac* | | *a* | *c* | *4ac* | *b* (*so that* $b^2 \geq 4ac$) | *No. of cases* |
|---|---|---|---|---|---|---|
| 1 | | 1 | 1 | 4 | 2, 3, 4, 5 | $1 \times 5 = 5$ |
| 2 | (*i*) | 1 | 2 | 8 | 3, 4, 5, 6 | $2 \times 4 = 8$ |
| | (*ii*) | 2 | 1 | | | |
| 3 | (*i*) | 1 | 3 | 12 | 4, 5, 6 | $2 \times 3 = 6$ |
| | (*ii*) | 3 | 1 | | | |
| 4 | (*i*) | 1 | 4 | 16 | 4, 5, 6 | $3 \times 3 = 9$ |
| | (*ii*) | 4 | 1 | | | |
| | (*iii*) | 2 | 2 | | | |
| 5 | (*i*) | 1 | 5 | 20 | 5, 6 | $2 \times 2 = 4$ |
| | (*ii*) | 5 | 1 | | | |
| 6 | (*i*) | 1 | 6 | 24 | 5, 6 | $4 \times 2 = 8$ |
| | (*ii*) | 6 | 1 | | | |
| | (*iii*) | 3 | 2 | | | |
| | (*iv*) | 2 | 3 | | | |
| 7 | ( *ac* = 7 *is not possible* ) | | | | | |
| 8 | (*i*) | 2 | 4 | 32 | 6 | $2 \times 1 = 2$ |
| | (*ii*) | 4 | 2 | | | |
| 9 | | 3 | 3 | 36 | 6 | 1 |
| | | | | | | *Total* = 43 |

Since $b^2 \geq 4ac$ and since the maximum value of $b^2$ is 36, $ac = 10, 11, 12, \ldots$ etc. is not possible.

Hence total number of favourable cases = 43.

$\therefore$ Required probability $= \dfrac{43}{216}$

**Example 4·9.** *The sum of two non-negative quantities is equal to $2n$. Find the chance that their product is not less than $\frac{3}{4}$ times their greatest product.*

**Solution.** Let $x > 0$ and $y > 0$ be the given quantities so that $x + y = 2n$.

We know that the product of two positive quantities whose sum is constant is greatest when the quantities are equal. Thus the product of $x$ and $y$ is maximum when $x = y = n$.

$\therefore$ Maximum product $= n \, . \, n = n^2$

Now $$P\left[xy \not< \frac{3}{4}n^2\right] = P\left[xy \geq \frac{3}{4}n^2\right] = P\left[x(2n-x) \geq \frac{3}{4}n^2\right]$$
$$= P\left[(4x^2 - 8nx + 3n^2) \leq 0\right]$$
$$= P\left[(2x - 3n)(2x - n) \leq 0\right]$$
$$= P\left[x \text{ lies between } \frac{n}{2} \text{ and } \frac{3n}{2}\right]$$

$\therefore$ Favourable range $= \dfrac{3n}{2} - \dfrac{n}{2} = n$

Total range $= 2n$

$\therefore$ Required probability $= \dfrac{n}{2n} = \dfrac{1}{2}$

**Example 4·10.** *Out of $(2n+1)$ tickets consecutively numbered three are drawn at random. Find the chance that the numbers on them are in A.P.*

**[Calicut Univ. B.Sc., 1991; Delhi Univ. B.Sc.(Stat. Hons.), 1992]**

**Solution.** Since out of $(2n+1)$ tickets, 3 tickets can be drawn in ${}^{2n+1}C_3$ ways,

$$\text{Exhaustive number of cases} = {}^{2n+1}C_3 = \frac{(2n+1)\,2n\,(2n-1)}{3\,!}$$
$$= \frac{n\,(4n^2-1)}{3}$$

To find the favourable number of cases we are to enumerate all the cases in which the numbers on the drawn tickets are in A.P with common difference, (say $d = 1, 2, 3, \ldots, n-1, n$).

If $d = 1$, the possible cases are as follows:

$$\left.\begin{array}{ccc} 1, & 2, & 3 \\ 2, & 3, & 4 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 2n-1, & n, & 2n+1 \end{array}\right\}, \text{ i.e., } (2n-1) \text{ cases in all}$$

If $d = 2$, the possible cases are as follows:

$$\left.\begin{array}{ccc} 1, & 3, & 5 \\ 2, & 4, & 6 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 2n-3, & 2n-1, & 2n+1 \end{array}\right\}, \text{ i.e., } (2n-3) \text{ cases in all}$$

and so on.

If $d = n-1$, the possible cases are as follows:

$$\left.\begin{array}{lll} 1, n, & 2n-1 \\ 2, n+1, & 2n \\ 3, n+2, & 2n+1 \end{array}\right\}, \textit{i.e.}, \text{ 3 cases in all}$$

If $d = n$, there is only one case, *viz.*, $(1, n+1, 2n+1)$.

Hence total number of favourable cases

$$= (2n-1) + (2n-3)) + ... + 5 + 3 + 1$$
$$= 1 + 3 + 5 + .... + (2n-1),$$

which is a series in A.P. with $d = 2$ and $n$ terms.

$\therefore$ Number of favourable cases $= \frac{n}{2}[1 + (2n-1)] = n^2$

$\therefore$ Required probability $= \dfrac{n^2}{n\ (4n^2-1)/3} = \dfrac{3n}{(4n^2-1)}$

## EXCERCISE 4 (a)

**1.** (*a*) Give the classical and statistical definitions of probability. What are the objections raised in these definitions?

**[Delhi Univ. B.Sc. (Stat. Hons.), 1988, 1985]**

(*b*) When are a number of cases said to be equally likely? Give an example each of the following :

(*i*) the equally likely cases,
(*ii*) four cases which are not equally likely, and
(*iii*) five cases in which one case is more likely than the other four.

(*c*) What is meant by mutually exclusive events? Give an example of

(*i*) three mutually exclusive events,
(*ii*) three events which are not mutually exclusive.

**[Meerut Univ. B.Sc. (Stat.), 1987]**

(*d*) Can

(*i*) events be mutually exclusive and exhaustive?
(*ii*) events be exhaustive and indepenent?
(*iii*) events be mutually exclusive and independent?
(*iv*) events be mutually exhaustive, exclusive and independent?

**2.** (*a*) Prove that the probability of obtaining a total of 9 in a single throw with two dice is one by nine.

(*b*) Prove that in a single throw with a pair of dice the probability of getting the sum of 7 is equal to 1/6 and the probability of getting the sum of 10 is equal to 1/12.

(*c*) Show that in a single throw with two dice, the chance of throwing more than seven is equal to that of throwing less than seven.

**Ans.** 5/12 **[Delhi Univ. B.Sc., 1987, 1985]**

(*d*) In a single throw with two dice, what is the number whose probability is minimum?

*(e)* Two persons $A$ and $B$ throw three dice (six faced). If $A$ throws 14, find $B$'s chance of throwing a higher number. **[Meerut Univ. B.Sc.(Stat.), 1987]**

3. *(a)* A bag contains 7 white, 6 red and 5 black balls. Two balls are drawn at random. Find the probability that they will both be white.

**Ans.** 21/153

*(b)* A bag contains 10 white, 6 red, 4 black and 7 blue balls. 5 balls are drawn at random. What is the probability that 2 of them are red and one black?

**Ans.** ${}^6C_2 \times {}^4C_1 / {}^{27}C_5$

4. *(a)* From a set of raffle tickets numbered 1 to 100, three are drawn at random. What is the probability that all the tickets are odd-numbered?

**Ans.** ${}^{50}C_3 / {}^{100}C_3$

*(b)* A number is chosen from each of the two sets :

$$(1, 2, 3, 4, 5, 6, 7, 8, 9); \quad (4, 5, 6, 7, 8, 9)$$

If $p_1$ is the probability that the sum of the two numbers be 10 and $p_2$ the probability that their sum be 8, find $p_1 + p_2$.

*(c)* Two different digits are chosen at random from the set 1,2,3,...,8. Show that the probability that the sum of the digits will be equal to 5 is the same as the probability that their sum will exceed 13, each being 1/14. Also show that the chance of both digits exceeding 5 is 3/28. **[Nagpur Univ. B.Sc., 1992]**

5. What is the chance that *(i)* a leap year selected at random will contain 53 Sundays? *(ii)* a non-leap year selected at random would contain 53 Sundays.

**Ans.** (i) 2/7, (ii) 1/7

6. *(a)* What is the probability of having a knave and a queen when two cards are drawn from a pack of 52 cards?

**Ans.** 8/663

*(b)* Seven cards are drawn at random from a pack of 52 cards. What is the probability that 4 will be red and 3 black?

**Ans.** ${}^{26}C_4 \times {}^{26}C_3 / {}^{52}C_7$

*(c)* A card is drawn from an ordinary pack and a gambler bets that it is a spade or an ace. What are the odds against his winning the bet?

**Ans.** 9:4

*(d)* Two cards are drawn from a pack of 52 cards. What is the chance that

*(i)* they belong to the same suit?

*(ii)* they belong to different suits and different denominations.

**[Bombay Univ. B.Sc., 1986]**

7. *(a)* If the letters of the word RANDOM be arranged at random, what is the chance that there are exactly two letters between A and O.

*(b)* Find the probability that in a random arrangement of the leters of the word 'UNIVERSITY', the two I's do not come together.

(c) In random arrangements of the letters of the word 'ENGINEERING', what is the probability that vowels always occur together?

**[Kurushetra Univ. B.E., 1991]**

(d) Letters are drawn one at a time from a box containing the letters A, H, M, O, S, T. What is the probability that the letters in the order drawn spell the word 'Thomas'?

**8.** A letter is taken out at random out of 'ASSISTANT' and a letter out of 'STATISTIC'. What is the chance that they are the same letters?

**9.** (a) Twelve persons amongst whom are $x$ and $y$ sit down at random at a round table. What is the probability that there are two persons between $x$ and $y$?

(b) $A$ and $B$ stand in a line at random with 10 other people. What is the probability that there will be 3 persons between $A$ and $B$?

**10.** (a) If from a lot of 30 tickets marked 1, 2, 3, ..., 30 four tickets are drawn, what is the chance that those marked 1 and 2 are among them?

**Ans.** 2/145

(b) A bag contains 50 tickets numbered 1, 2, 3, ..., 50 of which five are drawn at random and arranged in ascending order of the magnitude ($x_1 < x_2 < x_3 < x_4 < x_5$). What is the probability that $x_3 = 30$?

**Hint.** (a) Exhaustive number of cases = ${}^{30}C_4$

If, of the four tickets drawn, two tickets bear the numbers 1 and 2, the remaining 2 must have come out of 28 tickets numbered from 3 to 30 and this can be done in ${}^{28}C_2$ ways.

$\therefore$ Favourable number of cases = ${}^{28}C_2$

(b) Exhaustive number of cases = ${}^{50}C_5$

If $x_3 = 30$, then the two tickets with numbers $x_1$ and $x_2$ must have come out of 29 tickets numbered from 1 to 29 and this can be done in ${}^{29}C_2$ ways, and the other two tickets with numbers $x_4$ and $x_5$ must have been drawn out of 20 tickets numbered from 31 to 50 and this can be done in ${}^{20}C_2$ ways.

$\therefore$ No. of favourable cases = ${}^{29}C_2 \times {}^{20}C_2$.

**11.** Four persons are chosen ar random from a group containing 3 men, 2 women and 4 children. Show that the chance that exactly two of them will be children is 10/21. **[Delhi Univ. B.A.1988]**

$$\text{Ans. } \frac{{}^4C_2 \times {}^5C_2}{{}^9C_4} = \frac{10}{21}$$

**12.** From a group of 3 Indians, 4 Pakistanis and 5 Americans a sub-committee of four people is selected by lots. Find the probability that the sub-committee will consist of

(i) 2 Indians and 2 Pakistanis

(ii) 1 Indian, 1 Pakistani and 2 Americans *(iii)* 4 Americans [Madras Univ. B.Sc.(Main Stat.), 1987]

**Ans.** *(i)* $\dfrac{{}^3C_2 \times {}^4C_2}{{}^{12}C_4}$, *(ii)* $\dfrac{{}^3C_1 \times {}^4C_1 \times {}^5C_2}{{}^{12}C_4}$, *(iii)* $\dfrac{{}^5C_4}{{}^{12}C_4}$

**13.** In a box there are 4 granite stones, 5 sand stones and 6 bricks of identical size and shape. Out of them 3 are chosen at random. Find the chance that :

*(i)* They all belong to different varieties.

*(ii)* They all belong to the same variety.

*(iii)* They are all granite stones. **(Madras Univ. B.Sc., Oct. 1992)**

**14.** If $n$ people are seated at a round table, what is the chance that two named individuals will be next to each other?

**Ans.** $2/(n-1)$

**15.** Four tickets marked 00, 01, 10 and 11 respectively are placed in a bag. A ticket is drawn at random five times, being replaced each time. Find the probability that the sum of the numbers on tickets thus drawn is 23.

**[Delhi Univ. B.Sc.(Subs.), 1988]**

**16.** From a group of 25 persons, what is the probability that all 25 will have different birthdays? Assume a 365 day year and that all days are equally likely.

**[Delhi Univ. B.Sc.(Maths Hons.), 1987]**

**Hint.** $(365 \times 364 \times \ldots \times 341) \div (365)^{25}$

**4·4. Mathematical Tools : Preliminary Notions of Sets.** The set theory was developed by the German mathematician, G. Cantor (1845–1918).

**4·4·1. Sets and Elements of Sets.** A set is a well defined collection or aggregate of all possible objects having given properties and specified according to a well defined rule. The objects comprising a set are called elements, members or points of the set. Sets are often denoted by capital letters, *viz.*, $A, B, C$, etc. If $x$ is an element of the set $A$, we write symbolically $x \in A$ ($x$ belongs to $A$). If $x$ is not a member of the set $A$, we write $x \notin A$ ($x$ does not belong to $A$). Sets are often described by describing the properties possessed by their members. Thus the set $A$ of all non-negative rational numbers with square less than 2 will be written as $A = \{x : x \text{ rational}, x \geq 0, x^2 < 2\}$.

If every element of the set $A$ belongs to the set $B$, *i.e.*, if $x \in A \Rightarrow x \in B$, then we say that $A$ is a subset of $B$ and write symbolically $A \subseteq B$ ($A$ is contained in $B$) or $B \supseteq A$ ($B$ contains $A$). Two sets $A$ and $B$ are said to be *equal* or *identical* if $A \subseteq B$ and $B \subseteq A$ and we write $A = B$ or $B = A$.

A *null* or an *empty* set is one which does not contain any element at all and is denoted by $\phi$.

**Remarks. 1.** Every set is a subset of itself.

**2.** An empty set is subset of every set.

**3.** A set containing only one element is conceptually distinct from the element itself, but will be represented by the same symbol for the sake of convenience.

**4.** As will be the case in all our applications of set theory, especially to probability theory, we shall have a fixed set $S$ (say) given in advance, and we shall be concerned only with subsets of this given set. The underlying set $S$ may vary from one application to another, and it will be referred to as *universal set* of each particular discourse.

**4·4·2. Operation on Sets**

The union of two given sets $A$ and $B$, denoted by $A \cup B$, is defined as a set consisting of all those points which belong to either $A$ or $B$ or both. Thus symbolically,

$$A \cup B = \{ x : x \in A \text{ or } x \in B \}.$$

Similarly

$$\bigcup_{i=1}^{n} A_i = \{ x : x \in A_i \text{ for at least one } i = 1, 2, \ldots, n \}$$

The *intersection* of two sets $A$ and $B$, denoted by $A \cap B$, is defined as a set consisting of all those elements which belong to both $A$ and $B$. Thus

$$A \cap B = \{ x : x \in A \text{ and } x \in B \}.$$

Similarly

$$\bigcap_{i=1}^{n} A_i = \{ x : x \in A_i \text{ for all } i = 1, 2, \ldots, n \}$$

For example, if $A = \{1, 2, 5, 8, 10\}$ and $B = \{2, 4, 8, 12\}$, then

$A \cup B = \{1, 2, 4, 5, 8, 10, 12\}$ and $A \cap B = \{2, 8\}$.

If $A$ and $B$ have no common point, *i.e.*, $A \cap B = \phi$, then the sets $A$ and $B$ are said to be *disjoint*, *mutually exclusive* or *non-overlapping*.

The *relative difference* of a set $A$ from another set $B$, denoted by $A-B$ is defined as a set consisting of those elements of $A$ which do not belong to $B$. Symbolically,

$$A - B = \{ x : x \in A \text{ and } x \notin B \}.$$

The *complement* or *negative* of any set $A$, denoted by $\bar{A}$ is a set containing all elements of the universal set $S$, (say), that are not elements of $A$, *i.e.*, $\bar{A} = S - A$.

**4·4·3. Algebra of Sets**

Now we state certain important properties concerning operations on sets. If $A$, $B$ and $C$ are the subsets of a universal set $S$, then the following laws hold:

*Commutative Law* : $A \cup B = B \cup A$ , $A \cap B = B \cap A$

*Associative Law* : $(A \cup B) \cup C = A \cup (B \cup C)$

$(A \cap B) \cap C = A \cap (B \cap C)$

*Distributive Law* : $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

*Complementary Law* : $A \cup \bar{A} = S$ , $A \cap \bar{A} = \phi$

$A \cup S = S$ , $(\because A \subset S)$ , $A \cap S = A$

$A \cup \phi = A$ , $A \cap \phi = \phi$

*Difference Law* : $A - B = A \cap \bar{B}$

$A - B = A - (A \cap B) = (A \cup B) - B$

$A - (B - C) = (A - B) \cup (A - C)$·

$$(A \cup B) - C = (A - C) \cup (B - C)$$
$$A - (B \cup C) = (A - B) \cap (A - C)$$
$$(A \cap B) \cup (A - B) = A\,,\ (A \cap B) \cap (A - B) = \phi$$

*De-Morgan's Law—Dualization Law.*

$$\overline{(A \cup B)} = \overline{A} \cap \overline{B} \text{ and } \overline{(A \cap B)} = \overline{A} \cup \overline{B}$$

More generally

$$\overline{\left(\bigcup_{i=1}^{n} A_i\right)} = \bigcap_{i=1}^{n} \overline{A_i} \text{ and } \overline{\left(\bigcap_{i=1}^{n} A_i\right)} = \bigcup_{i=1}^{n} \overline{A_i}$$

*Involution Law* : $\overline{\overline{(A)}} = A$

*Idempotency Law* : $A \cup A = A\,,\ A \cap A = A$

**4·4·4. Limit of Sequence of Sets**

Let $\{A_n\}$ be a sequence of sets in $S$. The *limit supremum* or *limit superior* of the sequence, usually written as lim sup $A_n$, is the set of all those elements which belong to $A_n$ for infinitely many $n$. Thus

$$\limsup_{n \to \infty} A_n = \{\, x : x \in A_n \text{ for infinitely many } n \,\} \qquad \ldots(4\cdot3)$$

The set of all those elements which belong to $A_n$ for all but a finite number of $n$ is called *limit infinimum or limit inferior* of the sequence and is denoted by lim inf $A_n$. Thus

$$\liminf_{n \to \infty} A_n = \{\, x : x \in A_n \text{ for all but a finite number of } n \,\} \qquad \ldots(4\cdot3\ a)$$

The sequence $\{A_n\}$ is said to have a limit if and only if lim sup $A_n$ = lim inf $A_n$ and this common value gives the limit of the sequence.

**Theorem 4·1.** *lim sup* $A_n = \bigcap_{m=1}^{\infty} \left(\bigcup_{n=m}^{\infty} A_n\right)$

*and* *lim inf* $A_n = \bigcup_{m=1}^{\infty} \left(\bigcap_{n=m}^{\infty} A_n\right)$

**Def.** $\{A_n\}$ is a monotone (infinite) sequence of sets if either

*(i)* $A_n \subset A_{n+1}\ \forall\ n$ or *(ii)* $A_n \supset A_{n+1}\ \forall\ n$.

In the former case the sequence $\{A_n\}$ is said to be *non-decreasing sequence* and is usually expressed as $A_n \uparrow$ and in the latter case it is said to be *non-increasing sequence* and is expressed as $A_n \downarrow$.

For a monotone sequence (non-increasing or non-decreasing), the limit always exists and we have,

$$\lim_{n \to \infty} A_n = \begin{cases} \bigcup_{n=1}^{\infty} A_n \text{ in case } (i)\text{, i.e., } A_n \uparrow \\ \bigcap_{n=1}^{\infty} A_n \text{ in case } (ii)\text{, i.e., } A_n \downarrow \end{cases}$$

**4·4·5. Classes of Sets.** A group of sets will be termed as a *class* (of sets). Below we shall define some useful types of classes.

A ring **R** is a *non-empty* class of sets which is closed under the formation of 'finite unions' and 'difference',

*i.e.*, if $A \in \mathbf{R}$, $B \in \mathbf{R}$, then $A \cup B \in \mathbf{R}$ and $A - B \in \mathbf{R}$.

Obviously $\phi$ is a member of every ring.

A *field* **F** (or an *algebra*) is a non-empty class of sets which is closed under the formation of finite unions and under complementation. Thus

(i) $A \in \mathbf{F}, B \in \mathbf{F} \Rightarrow A \cup B \in \mathbf{F}$ and

(ii) $A \in \mathbf{F} \Rightarrow \bar{A} \in \mathbf{F}$.

A *σ-ring* **C** is a non-empty class of sets which is closed under the formation of 'countable unions' and 'difference'. Thus

(i) $A_i \in \mathbf{C},\ i = 1, 2, \ldots \quad \Rightarrow \quad \bigcup_{i=1}^{\infty} A_i \in \mathbf{C}$

(ii) $A \in \mathbf{C},\ B \in \mathbf{C} \quad \Rightarrow \quad A - B \in \mathbf{C}$.

More precisely σ-ring is a ring which is closed under the formation of countable unions.

A σ field (or σ-algebra) **B** is a non-empty class of sets that is closed under the formation of 'countable unions' and complementations,

*i.e.*,

(i) $A_i \in \mathbf{B},\ i = 1, 2, \ldots \quad \Rightarrow \quad \bigcup_{i=1}^{\infty} A_i \in \mathbf{B}$.

(ii) $A \in \mathbf{B} \quad \Rightarrow \quad \bar{A} \in \mathbf{B}$.

σ-field may also be defined as a field which is closed under the formation of countable unions.

**4·5. Axiomatic Approach to Probability.** The axiomatic approach to probability, which closely relates the theory of probability with the modern metric theory of functions and also set theory, was proposed by A.N. Kolmogorov, a Russian mathematician, in 1933. The axiomatic definition of probability includes 'both' the classical and the statistical definitions as particular cases and overcomes the deficiencies of each of them. On this basis, it is possible to construct a logically perfect structure of the modern theory of probability and at the same time to satisfy the enchanced requirements of modern natural science. The axiomatic development of mathematical theory of probability relies entirely upon the logic of deduction.

The diverse theorems of probability, as were available prior to 1933, were finally brought together into a unified axiomised system in 1933. It is important to remark that probability theory, in common with all axiomatic mathematical systems, is concerned solely with relations among undefined things.

The axioms thus provide a set of rules which define relationships between abstract entities. These rules can be used to deduce theorems, and the theorems can be brought together to deduce more complex theorems. These theorems have no empirical meaning although they can be given an interpretation in terms of empirical phenomenon. The important point, however, is that the mathematical development of probability theory is, in no way, conditional upon the interpretation given to the theory.

More precisely, under axiomatic approach, the probability can be deduced from mathematical concepts. To start with some concepts are laid down. Then some statements are made in respect of the properties possessed by these concepts. These properties, often termed as "*axioms*" of the theory, are used to frame some theorems. These theorems are framed without any reference to the real world and are deductions from the axioms of the theory.

**4·5·1. Random Experiment, Sample Space.** The theory of probability provides *mathematical models* for "real-world phenomenon" involving games of chance such as the tossing of coins and dice. We feel intuitively that statements such as

*(i)* "The probability of getting a "head" in one toss of an unbiased coin is 1/2"

*(ii)* "The probability of getting a "four" in a single toss of an unbiased die is 1/6",

should hold. We also feel that the probability of obtaining *either* a "5" or a "6" in a single throw of a die, should be the sum of the probabilities of a "5" and a "6", *viz.*, 1/6+1/6=1/3. That is, probabilities should have some kind of *additive* property. Finally, we feel that in a large number of repetitions of, for example, a coin tossing experiment, the proportion of heads should be approximately 1/2. That is, the probability should have a *frequency interpretation.*

To deal with these properties sensibly, we need a *mathematical description* or *model* for the probabilistic situation we have. Any such probabilistic situation is referred to as a *random experiment*, denoted by **E**. **E** may be a coin or die throwing experiment, drawing of cards from a well-shuffled pack of cards, an agricultural experiment to determine the effects of fertilizers on yield of a commodity, and so on.

Each performance in a random experiment is called a *trial*. That is, all the trials conducted under the same set of conditions form a random experiment. The result of a trial in a random experiment is called an outcome, an elementary event or a sample point. The totality of all possible outcomes (*i.e.*, sample points) of a random experiment constitutes the *sample space*.

Suppose $e_1, e_2, ..., e_n$ are the possible outcomes of a random experiment E such that no two or more of them can occur simultaneously and exactly one of the outcomes $e_1, e_2, ..., e_n$ must occur. More specifically, with an experiment E, we associated a set $S = (e_1, e_2, ..., e_n)$ of possible outcomes with the following properties:

*(i)* each element of $S$ denotes a possible outcome of the experiement,

*(ii)* any trial results in an outcome that corresponds to one and only one element of the set $S$.

The set $S$ associated with an experiment E, real or conceptual, satisfying the above two properties is called the *sample space* of the experiment.

**Remarks. 1.** The sample space serves as universal set for all questions concerned with the experiment.

2. A sample space $S$ is said to be finite (infinite) sample sapce if the number of elements in $S$ is finite (infinite). For example, the sample space associated with the experiment of throwing the coin until a head appears, is infinite, with possible *sample points*

$$\{\omega_1,\ \omega_2,\ \omega_3,\ \omega_4,\ ....\}$$

where $\omega_1 = H,\ \omega_2 = TH,\ \omega_3 = TTH,\ \omega_4 = TTTH$, and so on, $H$ denoting a head and $T$ a tail.

3. A sample space is called discrete if it contains only finitely or infinitely many points which can be arranged into a simple sequence $\omega_1,\ \omega_2,\ ...,$ while a sample space containing non-denumerable number of points is called a continuous sample space. In this book, we shall restrict ourselves to discrete sample spaces only.

**4·5·2. Event.** Every non-empty subset $A$ of $S$, which is a disjoint union of single element subsets of the sample space $S$ of a random experiment E is called an event. The notion of an event may also be defined as follows:

"Of all the possible outcomes in the sample space of an experiment,some outcomes satisfy a specified description,which we call an event."

**Remarks. 1.** As the empty set $\phi$ is a subset of $S$ , $\phi$ is also an event, known as *impossible* event.

2. An event $A$, in particular, can be a single element subset of $S$, in which case it is known as *elementary* event.

**4·5·3. Some Illustrations — Examples.** We discuss below some examples to illustrate the concepts of sample space and event.

1. Consider tossing of a coin singly. The possible outcomes for this experiment are (writing $H$ for a "head" and $T$ for a "tail") : $H$ and $T$. Thus the sample space $S$ consists of two points $\{\omega_1,\ \omega_2\}$, corresponding to each possible outcome or elementary event listed.

*i.e.,* $S = \{\omega_1,\ \omega_2\} = \{H,T\}$ and $n(S) = 2$,

where $n(S)$ is the total number of sample points in $S$.

If we consider two tosses of a coin, the possible outcomes are $HH, HT, TH, TT$. Thus, in this case the sample space S consists of four points $\{\omega_1, \omega_2, \omega_3, \omega_4\}$, corresponding to each possible outcome listed and $n(S) = 4$. Combinations of these outcomes form what we call events. For example, the event of getting at least one head is the set of the outcomes $\{HH, HT, TH\} = \{\omega_1, \omega_2, \omega_3\}$. Thus, mathematically, the events are subsets of $S$.

2. Let us consider a single toss of a die. Since there are six possible outcomes, our sample space $S$ is now a space of six points $\{\omega_1, \omega_2, \ldots, \omega_6\}$ where $\omega_i$ corresponds to the appearance of number $i$. Thus $S=\{\omega_1, \omega_2, \ldots, \omega_6\}=\{1,2,\ldots,6\}$ and $n(S)=6$. The event that the outcome is even is represented by the set of points $\{\omega_2, \omega_4, \omega_6\}$.

3. A coin and a die are tossed together. For this experiment, our sample space consists of twelve points $\{\omega_1, \omega_2, \ldots, \omega_{12}\}$ where $\omega_i$ $(i = 1, 2, \ldots, 6)$ represents a head on coin together with appearance of $i$th number on the die and $\omega_i$ $(i = 7, 8, \ldots, 12)$ represents a tail on coin together with the appearance of $i$th number on die. Thus

$$S = \{\omega_1, \omega_2, \ldots, \omega_{12}\} = \{(H, T) \times (1, 2, \ldots, 6)\} \text{ and } n(S) = 12$$

**Remark.** If the coin and die are unbiased, we can see intuitively that in each of the above examples, the outcomes (sample points) are equally likely to occur.

4. Consider an experiment in which two balls are drawn one by one from an urn containing 2 white and 4 blue balls such that when the second ball is drawn, the first is *not* replaced.

Let us number the six balls as 1, 2, 3, 4, 5 and 6, numbers 1 and 2 representing a white ball and numbers 3, 4, 5, and 6 representing a blue ball. Suppose in a draw we pick up balls numbered 2 and 6. Then (2,6) is called an outcome of the experiment. It should be noted that the outcome (2,6) is different from the outcome (6,2) because in the former case ball No. 2 is drawn first and ball No.6 is drawn next while in the latter case, 6th ball is drawn first and the second ball is drawn next.

The sample space consists of thirty points as listed below:

| | | | | |
|---|---|---|---|---|
| $\omega_1 = (1,2)$ | $\omega_2 = (1,3)$ | $\omega_3 = (1,4)$ | $\omega_4 = (1,5)$ | $\omega_5 = (1,6)$ |
| $\omega_6 = (2,1)$ | $\omega_7 = (2,3)$ | $\omega_8 = (2,4)$ | $\omega_9 = (2,5)$ | $\omega_{10} = (2,6)$ |
| $\omega_{11} = (3,1)$ | $\omega_{12} = (3,2)$ | $\omega_{13} = (3,4)$ | $\omega_{14} = (3,5)$ | $\omega_{15} = (3,6)$ |
| $\omega_{16} = (4,1)$ | $\omega_{17} = (4,2)$ | $\omega_{18} = (4,3)$ | $\omega_{19} = (4,5)$ | $\omega_{20} = (4,6)$. |
| $\omega_{21} = (5,1)$ | $\omega_{22} = (5,2)$ | $\omega_{23} = (5,3)$ | $\omega_{24} = (5,4)$ | $\omega_{25} = (5,6)$ |
| $\omega_{26} = (6,1)$ | $\omega_{27} = (6,2)$ | $\omega_{28} = (6,3)$ | $\omega_{29} = (6,4)$ | $\omega_{30} = (6,5)$ |

Thus

$$S = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_{30}\} \text{ and } n(S) = 30$$

$$\Rightarrow \quad S = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\} - \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$$

The event

*(i)* the first ball drawn is white

*(ii)* the second ball drawn is white

*(iii)* both the balls drawn are white

*(iv)* both the balls drawn are black

are represented respectively by the following sets of points:

$\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9, \omega_{10}\}$,

$\{\omega_1, \omega_6, \omega_{11}, \omega_{12}, \omega_{16}, \omega_{17}, \omega_{21}, \omega_{22}, \omega_{26}, \omega_{27}\}$,

$\{\omega_1, \omega_6\}$, and

$\{\omega_{13}, \omega_{14}, \omega_{15}, \omega_{18}, \omega_{19}, \omega_{20}, \omega_{23}, \omega_{24}, \omega_{25}, \omega_{28}, \omega_{29}, \omega_{30}\}$.

5. Consider an experiment in which two dice are tossed. The sample space $S$ for this experiment is given by

$$S = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$$

and $\quad n(S) = 6 \times 6 = 36.$

Let $E_1$ be the event that 'the sum of the spots on the dice is greater than 12', $E_2$ be the event that 'the sum of spots on the dice is divisible by 3', and $E_3$ be the event that 'the sum is greater than or equal to two and is less than or equal to 12'. Then these events are represented by the following subsets of $S$ :

$$E_1 = \{\phi\},\ E_3 = S \text{ and}$$
$$E_2 = \{(1, 2), (1, 5), (2, 1), (2, 4), (3, 3), (3, 6), (4, 2), (4, 5), (5, 1), (5, 4), (6, 3), (6, 6)\}$$

Thus $\quad n(E_1)=0,\ n(E_2)=12,$ and $n(E_3)=36$

Here $E$ is an 'impossible event' and $E_3$ a 'certain event'.

6. Let E denote the experiment of tossing a coin three times in succession or tossing three coins at a time. Then the sample space $S$ is given by

$$\begin{aligned} S &= \{H, T\} \times \{H, T\} \times \{H, T\} \\ &= \{H, T\} \times \{HH, HT, TH, TT\} \\ &= \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\} \\ &= \{\omega_1, \omega_2, \omega_3, \ldots, \omega_8\}, \text{ say.} \end{aligned}$$

If $E_1$ is the event that 'the number of heads exceeds the number of tails', $E_2$, the event of 'getting two heads' and $E_3$, the event of getting 'head in the first trial' then these are represented by the following sets of points :

$$E_1 = \{\omega_1, \omega_2, \omega_3, \omega_5\},$$
$$E_2 = \{\omega_2, \omega_3, \omega_5\}$$
and $\quad E_3 = \{\omega_1, \omega_2, \omega_3, \omega_4\}.$

7. In the foregoing examples the sample sapce is finite. To construct an experiment in which the sample sapce is countably infinite, we toss a coin repeatedly until head or tail appears twice in succession. The sample space of all the possible outcomes may be represented as :

$$S = \{HH, TT, THH, HTT, HTHH, THTT, THTHH, HTHTT, \ldots\}.$$

**4·5·4. Algebra of Events.** For events $A, B, C$

(i) $A \cup B = \{\omega \in S : \omega \in A \text{ or } \omega \in B\}$

(ii) $A \cap B = \{\omega \in S : \omega \in A \text{ and } \omega \in B\}$

(iii) $\bar{A}$ (*A complement*) $= \{\omega \in S : \omega \notin A\}$

(iv) $A - B = \{\omega \in S : \omega \in A \text{ but } \omega \notin B\}$

(v) Similar generalisations for $\bigcup_{i=1}^{n} A_i,\ \bigcap_{i=1}^{n} A_i,\ \bigcup_{i} A_i$ etc.

(vi) $A \subset B \Rightarrow$ for every $\omega \in A,\ \omega \in B$.

(vii) $B \supset A \Rightarrow A \subset B$.

(viii) $A = B$ if and only if $A$ and $B$ have the same elements, *i.e.*, if $A \subset B$ and $B \subset A$.

(ix) $A$ and $B$ disjoint ( mutually exclusive ) $\Rightarrow A \cap B = \phi$ (*null set*).

(x) $A \cup B$ can be denoted by $A + B$ if $A$ and $B$ are disjoint.

(xi) $A \Delta B$ denotes those $\omega$ belonging to exactly one of $A$ and $B$, *i.e.*,

$$A \Delta B = A\overline{B} \cup \overline{A}B$$

**Remark.** Since the events are subsets of $S$, all the laws of set theory *viz.*, commutative laws, associative laws, distributive laws, De-Morgan's law, etc., hold for algebra of events.

**Table – *Glossary of Probability Terms***

| Statement | Meaning in terms of set theory |
|---|---|
| 1. At least one of the events $A$ or $B$ occurs. | $\omega \in A \cup B$ |
| 2. Both the events $A$ and $B$ occur. | $\omega \in A \cap B$ |
| 3. Neither $A$ nor $B$ occurs | $\omega \in \overline{A} \cap \overline{B}$ |
| 4. Event $A$ occurs and $B$ does not occur | $\omega \in A \cap \overline{B}$ |
| 5. Exactly one of the events $A$ or $B$ occurs. | $\omega \in A \Delta B$ |
| 6. Not more than one of the events $A$ or $B$ occurs. | $\omega \in (A \cap \overline{B}) \cup (\overline{A} \cap B) \cup (\overline{A} \cap \overline{B})$ |
| 7. If event $A$ occurs, so does $B$ | $A \subset B$ |
| 8. Events $A$ and $B$ are mutually exclusive. | $A \cap B = \phi$ |
| 9. Complementary event of $A$. | $\overline{A}$ |
| 10. Sample space | universal set $S$ |

**Example 4·11.** *$A$, $B$ and $C$ are three arbitrary events. Find expressions for the events noted below, in the context of $A$, $B$ and $C$.*

(i) *only $A$ occurs,*

(ii) *Both $A$ and $B$, but not $C$, occur,*

(iii) *All three events occur,*

(iv) *At least one occurs,*

(v) *At least two occur,*

(vi) *One and no more occurs,*

(vii) *Two and no more occur,*

(viii) *None occurs.*

**Solution.**

(i) $A \cap \overline{B} \cap \overline{C}$, (ii) $A \cap B \cap \overline{C}$, (iii) $A \cap B \cap C$,

(iv) $A \cup B \cup C$,

(v) $(A \cap B \cap \bar{C}) \cup (A \cap \bar{B} \cap C) \cup (\bar{A} \cap B \cap C) \cup (A \cap B \cap C)$

(vi) $(A \cap \bar{B} \cap \bar{C}) \cup (\bar{A} \cap B \cap \bar{C}) \cup (\bar{A} \cap \bar{B} \cap C)$

(vii) $(A \cap B \cap \bar{C}) \cup (\bar{A} \cap B \cap C) \cup (A \cap \bar{B} \cap C)$

(viii) $\bar{A} \cap \bar{B} \cap \bar{C}$ or $\overline{A \cup B \cup C}$

## EXERCISE 4(b)

1. *(i)* If $A, B$ and $C$ are any three events, write down the theoretical expressions for the following events:

*(a)* Only $A$ occurs, *(b)* $A$ and $B$ occur but $C$ does not,

*(c)* $A, B$, and $C$ all the three occur, *(d)* at least one occurs

*(e)* At least two occur, *(f)* one does not occur,

*(g)* Two do not occurs, and *(h)* None occurs.

*(ii)* $A$, $B$ and $C$ are three events. Express the following events in appropriate symbols:

*(a)* Simultaneous occurrence of $A, B$ and $C$.

*(b)* C ·currence of at least one of them.

*(c)* $A$, $B$ and $C$ are mutually exclusive events.

*(d)* Every point of $A$ is contained in $B$.

*(e)* The event $B$ but not $A$ occurs. **[Gauhati Univ. B.Sc., Oct.1990]**

2. A sample space $S$ contains four points $x_1, x_2, x_3$ and $x_4$ and the values of a set function $P(A)$ are known for the following sets :

$$A_1 = (x_1, x_2) \text{ and } P(A_1) = \frac{4}{10};\ A_2 = (x_3, x_4) \text{ and } P(A_2) = \frac{6}{10};$$

$$A_3 = (x_1, x_2, x_3) \text{ and } P(A_3) = \frac{4}{10};\ A_4 = (x_2, x_3, x_4) \text{ and } P(A_4) = \frac{7}{10}$$

Show that :

*(i)* the total number of sets (including the "null" set of number points) of points of $x$ is 16.

*(ii)* Although the set containing no sample point has zero probability, the converse is not always true, *i.e.*, a set may have zero probability and yet it may be the set of a number of points.

3. Describe explicitly the sample spaces for each of the following experiments:

*(i)* The tossing of four coins.

*(ii)* The throwing of three dice.

*(iii)* The tossing of ten coins with the aim of observing the numbers of tails coming up.

*(iv)* Two cards are selected from a standard deck of cards.

*(v)* Four successive draws *(a)* with replacement, and *(b)* without replacement, from a bag containing fifty coloured balls out of which ten are white, twenty blue and twenty red.

*(vi)* A survey of families with two children is conduced and the sex of the children (the older child first) is recorded.

*(vii)* A survey of families with three children is made and the sex of the children (in order of age, oldest child first) are recorded.

*(viii)* Three distinguishable objects are distributed in three numbered cells.

*(ix)* A poker hand (five cards) is dealt from an ordinary deck of cards.

*(x)* Selecting $r$ screws from the lot produced by a machine, a screw can be defective or non-defective.

**4.** In an experiment a coin is thrown five times. Write down the sample space. How many points are there in the sample space?

**5.** Describe sample space appropriate in each of the following cases :

*(i)* $n$-tosses of a coin with head or tails as outcome in each toss.

*(ii)* Successive tosses of a coin until a head turns up.

*(iii)* A survey of families with two children is conducted and the sex of the children (the older child first) is recorded.

*(iv)* Two successive draws, *(a)* with replacement *(b)* without replacement, from a bag containing 4 coloured toys out of which one is white, one black and 2 red toys. **[M.S.Baroda Univ. B.Sc., 1991]**

**6.** *(a)* An experiment consists of tossing an unbiased coin until the same result appears twice on succession for the first time. To every possible outcome requiring $n$ tosses attribute probability $1/2^n$. Describe the sample space.

*(b)* A coin is tossed until there are either two consecutive heads or two consecutive tails or the number of tosses becomes five. Describe the sample space along with the probability associated with each sample point, if every sequence of $n$ tosses has probabilty $2^{-n}$. **[Civil Services (main), 1983]**

**7.** Urn 1 contains two white, one red and 3 black balls. Urn 2 contains one white, 3 red and 2 black balls. An experiment consists of first selecting an urn and then drawing a ball from this urn. Define a suitable sample space for this experiment.

**8.** Suppose an experiment has $n$ outcomes $A_1, A_2, ..., A_n$ and that it is repeated $r$ times. Let $x_1, x_2, ..., x_n$ record the number of occurrences of $A_1, A_2, ..., A_n$. Describe the sample space. Show that the number of sample points is

$$\binom{n+r-1}{r-1}$$

**9.** A manufacturer buys parts from four different vendors numbered 1, 2, 3 and 4. Referring to orders placed on two successive days, (1,4) denotes the event that on the first day, the order was given to vendor 1 and on the second day it was given to vendor 4. Letting $A$ represent the event that vendor 1 gets at least one of these two orders, $B$ the event that the same vendor gets both orders and $C$ the event that vendors 1 and 3 do not get either order. List the elements of :

*(a)* entire sample space, *(b)* $A$, *(c)* $B$, *(d)* $C$, *(e)* $\bar{A}$, *(f)* $\bar{B}$, *(g)* $B \cup C$, *(h)* $A \cap B$, *(i)* $A \cap C$, *(j)* $\overline{A \cup B}$, and *(k)* $A - B$

[**Hint.** *(a)* The elements of entire sample space are

(1,1); (1, 2); (1, 3); (1, 4); (2, 1); (2, 2); (2, 3); (2, 4);
(3, 1); (3, 2); (3, 3); (3, 4); (4, 1); (4, 2); (4, 3); (4, 4).

(b) The elements of $A$ are

$(1, 1); (1, 2); (1, 3); (1, 4); (2, 1); (3, 1); (4, 1)$.

(c) The elements of $B$ are $(1, 1); (2, 2); (3, 3)$ and $(4, 4)$.

(d) The elements of $C$ are $(2, 2); (2, 4); (4, 2); (4, 4)$.

(e) The element of $\overline{A}$ are :

$(2, 2); (2, 3); (2, 4); (3, 2); (3, 3); (3, 4); (4, 2); (4, 3); (4, 4)$.

(f) The elements of $\overline{B}$ are : $(1, 2); (1, 3); (1, 4); (2, 1); (2, 3); (2, 4); (3, 1); (3, 2); (3, 4); (4, 1); (4, 2); (4, 3)$.

(g) The elements of $B \cup C$ are $(1, 1); (2, 2); (3, 3); (4, 4); (2, 4); (4, 2)$.

(h) The elements of $A \cap B$ are $(1, 1)$.

(i) $A \cap C = \phi$

(j) Since $\overline{A \cup B} = \overline{A} \cap \overline{B}$. The elements of $\overline{A \cup B}$ are $(2, 3); (2, 4); (3, 2); (3, 4); (4, 2); (4, 3)$.

(k) The elements of $A - B$ are $(1, 2); (1, 3); (1, 4); (2, 1); (3, 1); (4, 1)$.

**4·6. Probability — Mathematical Notion.** We are now set to give the mathematical notion of the occurrence of a random phenomenon and the mathematical notion of probability. Suppose in a large number of trials the sample space $S$ contains $N$ sample points. The event $A$ is defined by a description which is satisfied by $N_A$ of the occurrences. The frequency interpretation of the probability $P(A)$ of the event $A$, tells us that $P(A)=N_A/N$.

A purely mathematical definition of probability cannot give us the actual value of $P(A)$ and this must be considered as a function defined on all events. With this in view, a mathematical definition of probability is enunciated as follows:

*"Given a sample description space, probability is a function which assigns a non-negative real number to every event A, denoted by P(A) and is called the probability of the event A."*

**4·6·1. Probability Function.** $P(A)$ is the probability function defined on a $\sigma$-field **B** of events if the following properties or axioms hold :

1. For each $A \in$ **B**, $P(A)$ is defined, is real and $P(A) \geq 0$

2. $P(S) = 1$

3. If $\{A_n\}$ is any finite or infinite sequence of disjoint events in **B**, then

$$P\left(\bigcup_{i=1}^{n} A_n\right) = \sum_{i=1}^{n} P(A_i) \qquad \text{...(4·4)}$$

The above three axioms are termed as the axiom of positiveness, certainty and union (additivity), respectively.

**Remarks. 1.** The set function **P** defined on $\sigma$-field **B**, taking its values in the real line and satisfying the above three axioms is called the probability measure.

**2.** The same definition of probability applies to *uncountable sample space* except that special restrictions must be placed on $S$ and its subsets. It is important to realise that for a complete description of a probability measure, three things must be specified, *viz.*, the sample space $S$, the $\sigma$-field ($\sigma$-algebra) **B** formed from certain subset of $S$ and set function **P**. The triplet (**S**, **B**, **P**) is often called the *probability space*. In most elementary applications, $S$ is finite and the $\sigma$-algebra **B** is taken to be the collection of all subsets of $S$.

3. It is interesting to see that there are some formal statements of the properties of events derived from the frequency approach. Since $P(A)=N_A/N$, it is easy to see that $P(A) \geq 0$, as in Axiom 1. Next since $N_S = N$, $P(S)=1$, as in Axiom 2. In case of two mutually exclusive (or disjoint) events $A$ and $B$ defined by sample points $N_A$ and $N_B$, the sample points belonging to $A \cup B$ are $N_A + N_B$. Therefore,

$$P(A \cup B) = \frac{N_A + N_B}{N} = \frac{N_A}{N} + \frac{N_B}{N} = P(A) + P(B), \text{ as in axiom 3.}$$

**Extended Axiom of Addition.** If an event $A$ can materialise in the occurrence of any one of the pairwise disjoint events $A_1$, $A_2$, ... so that

$$A = \bigcup_{i=1}^{\infty} A_i\ ;\ A_i \cap A_j = \phi \quad (i \neq j)$$

then

$$P(A) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \qquad \text{...(1)}$$

**Axiom of Continuity.** If $B_1, B_2, \ldots, B_n, \ldots$ be a countable sequences of events such that

$$(i)\ B_i \supset B_{i+1},\ (i = 1, 2, 3, \ldots)$$

and

$$(ii)\ \bigcap_{n=1}^{\infty} B_n = \phi$$

*i.e.*, if each succeeding event implies the preceeding event and if their simultaneous occurrence is an impossible event then

$$\lim_{n \to \infty} P(B_n) = 0 \qquad \text{...(2)}$$

We shall now prove that these two axioms, *viz.*, the extended axiom of addition and axiom of continuity are equivalent, *i.e.*, each implies the other, *i.e.*, (1) $\Leftrightarrow$ (2).

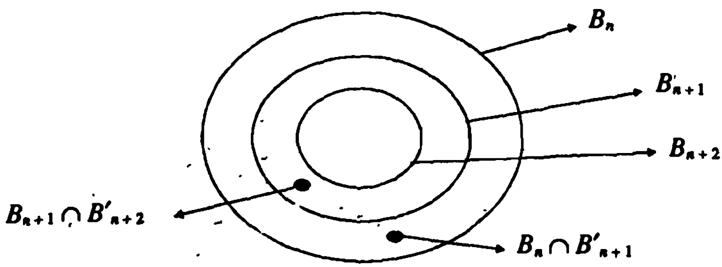
**Theorem 4·1.** *Axiom of continuity follows from the extended axiom of addition and vice versa.*

**Proof.** *(a)* (1) $\Rightarrow$ (2). Let $\{B_n\}$ be a countable sequence of events such that

$$B_1 \supset B_2 \supset B_3 \supset \ldots, \supset B_n \supset B_{n+1} \supset \ldots$$

and let for any $n \geq 1$,

$$\bigcap_{k \geq n} B_k = \phi \qquad (*)$$




Then it is obvious from the diagram that

$$B_n = B_n B'_{n+1} \cup B_{n+1} B'_{n+2} \cup \ldots \cup \left( \bigcap_{k \geq n} B_k \right)$$

$$\Rightarrow \qquad B_n = \left( \bigcup_{k=n}^{\infty} B_k B'_{k+1} \cup \left( \bigcap_{k \geq n} B_k \right) \right),$$

where the events $B_k B'_{k+1}$; $(k=n, n+1, \ldots)$ are pairwise disjoint and each is disjoint with $\bigcap_{k \geq n} B_k$.

Thus $B_n$ has been expressed as the countable union of pairwise disjoint events and hence by the extended axiom of addition, we get

$$P(B_n) = \sum_{k=n}^{\infty} P(B_k B'_{k+1}) + P\left( \bigcap_{k \geq n} B_k \right)$$

$$= \sum_{k=n}^{\infty} P(B_k B'_{k+1}), \qquad (**)$$

since, from (*)

$$P\left( \bigcap_{k \geq n} B_k \right) = P(\phi) = 0$$

Further, from (**), since

$$\sum_{k=1}^{\infty} P(B_k B'_{k+1}) = P(B_1) \leq 1,$$

the right hand sum in (**), being the remainder after $n$ terms of a convergent series tends to zero as $n \to \infty$.

Hence

$$\lim_{n \to \infty} P(B_n) = \lim_{n \to \infty} \sum_{k=n}^{\infty} P(B_k B'_{k+1}) = 0$$

Thus the extended axiom of addition implies the axiom of continuity.

*(b)* **Conversely (2) $\Rightarrow$ (1)**, *i.e., the extended axiom of addition follows from the axiom of continuity.*

Let $\{A_n\}$ be a countable sequence of pairwise disjoint events and let

$$A = \bigcup_{i=1}^{\infty} A_i$$

$$= \left(\bigcup_{i=1}^{n} A_i\right) \cup \left(\bigcup_{i=n+1}^{\infty} A_i\right) \qquad ...(3)$$

Let us define a countable sequence $\{B_n\}$ of events by

$$B_n = \bigcup_{i=n}^{\infty} A_i \qquad ...(4)$$

Obviously $B_n$ is a decreasing sequence of events, *i.e.*,

$$B_1 \supset B_2 \supset \ldots \supset B_n \supset B_{n+1} \supset \ldots. \qquad ...(5)$$

Also we have

$$A = \left(\bigcup_{i=1}^{n} A_i\right) \cup B_{n+1} \qquad ...(6)$$

Since $A_i$ 's are pairwise disjoint, we get

$$A_i \cap B_{n+1} = \phi, \ (i = 1, 2, ..., n) \qquad ...(6a)$$

From (4) we see that if the event $B_n$ has occurred it implies the occurrence of any one of the events $A_{n+1}, A_{n+2}, \ldots$ Without loss of generality let us assume that this event is $A_i$ $(i = n + 1, n + 2, \ldots)$. Further since $A_i$ 's are pairwise disjoint, the occurrence of $A_i$ implies that events $A_{i+1}, A_{i+2}, \ldots$ do not occur leading to the conclusion that $B_{i+1}, B_{i+2}, \ldots$ will not occur.

$$\Rightarrow \qquad \bigcap_{i=n}^{\infty} B_i = \phi \qquad ...(7)$$

From (5) and (7), we observe that both the conditions of axiom of continuity are satisfied and hence we get

$$\lim_{n\to\infty} P(B_n) = 0 \qquad ...(8)$$

From (6), we get

$$P(A) = P\left[\left(\bigcup_{i=1}^{n} A_i\right) \cup B_{n+1}\right]$$

$$= \sum_{i=1}^{n} P(A_i) + P(B_{n+1})$$

(By axiom of Additivity)

$$\Rightarrow \qquad P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i) + \lim_{n\to\infty} (B_{n+1})$$

$$= \sum_{i=1}^{\infty} P(A_i), \qquad \text{[From (8)]}$$

which is the extended axiom of addition.

## THEOREMS ON PROBABILITIES OF EVENTS

**Theorem 4·2.** *Probability of the impossible event is zero, i.e., $P(\phi) = 0$.*

**Proof.** Impossible event contains no sample point and hence the certain event $S$ and the impossible event $\phi$ are mutually exclusive.

Hence $\qquad S \cup \phi = S$

$\therefore \qquad P(S \cup \phi) = P(S)$

$\Rightarrow \qquad P(S) + P(\phi) = P(S)$ [By Axiom 3]

$\Rightarrow \qquad P(\phi) = 0$

**Remark.** It may be noted $P(A)=0$, does not imply that $A$ is necessarily an empty set. In practice, probability '0' is assigned to the events which are so rare that they happen only once in a lifetime. For example, if a person who does not know typing is asked to type the manuscript of a book, the probability of the event that he will type it correctly without any mistake is 0.

As another illustration, let us consider the random tossing of a coin. The event that the coin will stand erect on its edge, is assigned the probability 0.

The study of continuous random variable provides another illustration to the fact that $P(A)=0$, does not imply $A=\phi$, because in case of continous random variable $X$, the proability at a point is always zero, *i.e.*, $P(X=c)=0$ [See Chapter 5].

**Theorem 4·3.** Probability of the complementary event $\bar{A}$ of $A$ is given by

$$P(\bar{A}) = 1 - P(A)$$

**Proof.** $A$ and $\bar{A}$ are disjoint events.

Moreover, $\qquad A \cup \bar{A} = S$

From axioms 2 and 3 of probability, we have

$$P(A \cup \bar{A}) = P(A) + P(\bar{A}) = P(S) = 1$$

$\Rightarrow \qquad P(\bar{A}) = 1 - P(A)$

**Cor. 1.** We have $\qquad P(A) = 1 - P(\bar{A})$

$\Rightarrow \qquad P(A) \leq 1 \qquad (\because P(\bar{A}) \geq 0)$

**Cor. 2.** $P(\phi) = 0$, since $\phi = \bar{S}$

and $\qquad P(\phi) = P(\bar{S}) = 1 - P(S) = 1 - 1 = 0.$

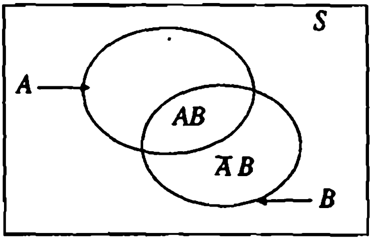
**Theorem 4·4.** *For any two events A and B,*

$$P(\bar{A} \cap B) = P(B) - P(A \cap B) \qquad \textbf{[Mysore Univ. B.Sc., 1992]}$$

**Proof.**

$\bar{A} \cap B$ and $A \cap B$ are disjoint events and

$(A \cap B) \cup (\bar{A} \cap B) = B$




Hence by axiom 3, we get

$$P(B) = P(A \cap B) + P(\bar{A} \cap B)$$

$$\Rightarrow P(\bar{A} \cap B) = P(B) - P(A \cap B)$$

**Remark.** Similarly, we shall get

$$P(A \cap \bar{B}) = P(A) - P(A \cap B)$$

**Theorem 4·5.** *Probability of the union of any two events A and B is given by*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

**Proof.** $A \cup B$ can be written as the union of the two mutually disjoint events, $A$ and $B \cap \bar{A}$.

$$\therefore \quad P(A \cup B) = P[A \cup (B \cap \bar{A})] = P(A) + P(B \cap \bar{A})$$

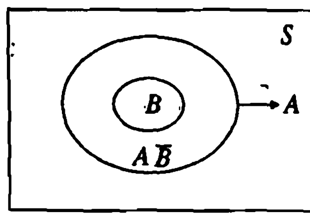
$$= P(A) + P(B) - P(A \cap B) \qquad \text{(c.f. Theorem 4·4)}$$

**Theorem 4·6.** *If $B \subset A$, then*

*(i)* $P(A \cap \bar{B}) = P(A) - P(B)$,

*(ii)* $P(B) \le P(A)$

**Proof.** *(i)* When $B \subset A$, $B$ and $A \cap \bar{B}$ are mutually exclusive events and their union is $A$




Therefore

$$P(A) = P[B \cup (A \cap \bar{B})]$$

$$= P(B) + P(A \cap \bar{B}) \qquad \text{[By axiom 3]}$$

$$\Rightarrow \quad P(A \cap \bar{B}) = P(A) - P(B)$$

*(ii)* Using axiom 1,

$$P(A \cap \bar{B}) \ge 0 \quad \Rightarrow \quad P(A) - P(B) \ge 0$$

Hence $\quad P(B) \le P(A)$

**Cor.** Since $(A \cap B) \subset A$ and $(A \cap B) \subset B$,

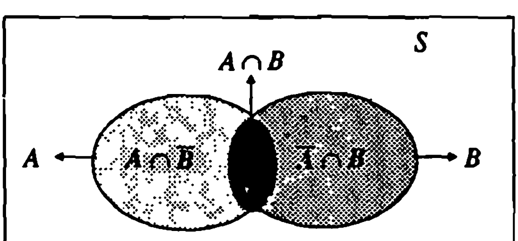
$$P(A \cap B) \le P(A) \quad \text{and} \quad P(A \cap B) \le P(B)$$

**4·6·2. Law of Addition of Probabilities**

**Statement.** *If A and B are any two events [subsets of sample space S] and are not disjoint, then*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad \text{...(4·5)}$$

**Proof.**



We have

$$A \cup B = A \cup (\bar{A} \cap B)$$

Since $A$ and $(\bar{A} \cap B)$ are disjoint,

$$P(A \cup B) = P(A) + P(\bar{A} \cap B)$$
$$= P(A) + [P(\bar{A} \cap B) + P(A \cap B)] - P(A \cap B)$$
$$= P(A) + P[(\bar{A} \cap B) \cup (A \cap B)] - P(A \cap B)$$
$$[\because (\bar{A} \cap B) \text{ and } (A \cap B) \textit{ are disjoint}]$$

$$\Rightarrow \quad P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

**Remark.** An alternative proof is provided by Theorems 4·4 *and* 4·5.

**4·6·3. Extention of General Law of Addition of Probabilities.** *For n events $A_1, A_2, \ldots, A_n$, we have*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum\sum_{1 \le i < j \le n} P(A_i \cap A_j) + \sum\sum\sum_{1 \le i < j < k \le n} P(A_i \cap A_j \cap A_k)$$
$$- \ldots + (-1)^{n-1} P(A_1 \cap A_2 \cap \ldots \cap A_n) \qquad \ldots(4\text{·}6)$$

**Proof.** For two events $A_1$ and $A_2$, we have

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \qquad \ldots(*)$$

Hence (4·6) is true for $n = 2$.

Let us now suppose that (4·6) is true for $n = r$, (say). Then

$$P\left(\bigcup_{i=1}^{r} A_i\right) = \sum_{i=1}^{r} P(A_i) - \sum\sum_{1 \le i < j \le r} P(A_i \cap A_j) + \ldots + (-1)^{r-1} P(A_1 \cap A_2 \cap \ldots \cap A_r) \qquad \ldots(**)$$

Now

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) = P\left[\left(\bigcup_{i=1}^{r} A_i\right) \cup A_{r+1}\right]$$

$$= P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) - P\left[\left(\bigcup_{i=1}^{r} A_i\right) \cap A_{r+1}\right] \qquad \ldots[\text{Using } (*)]$$

$$= P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) - P\left[\bigcup_{i=1}^{r} (A_i \cap A_{r+1})\right] \qquad \text{(Distributive Law)}$$

$$= \sum_{i=1}^{r} P(A_i) - \sum\sum_{1 \le i < j \le r} P(A_i \cap A_j) + \ldots$$
$$\ldots + (-1)^{r-1} P(A_1 \cap A_2 \cap \ldots \cap A_r) + P(A_{r+1})$$
$$- P\left[\bigcup_{i=1}^{r} (A_i \cap A_{r+1})\right] \qquad \ldots[\text{From } (**)]$$

$$= \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1 \le i < j \le r} P(A_i \cap A_j) + \ldots$$
$$+ (-1)^{r-1} P(A_1 \cap A_2 \cap \ldots \cap A_r)$$

$$- [\sum_{i=1}^{r} P(A_i \cap A_{r+1}) - \sum\sum_{1 \le i < j \le r} P(A_i \cap A_j \cap A_{r+1})$$

$$+ ... + (-1)^{r-1} P(A_1 \cap A_2 \cap ... \cap A_r \cap A_{r+1})] \quad ...[\text{From } (**)]$$

$$\Rightarrow \quad P(\bigcup_{i=1}^{r+1} A_i) = \sum_{i=1}^{r+1} P(A_i) - [\sum\sum_{1 \le i < j \le r} P(A_i \cap A_j) + \sum_{i=1}^{r} P(A_i \cap A_{r+1})]$$

$$+ ... + (-1)^r P[(A_1 \cap A_2 \cap ... \cap A_{r+1})]$$

$$= \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1 \le i < j \le (r+1)} P(A_i \cap A_j)$$

$$+ ... + (-1)^r P(A_1 \cap A_2 \cap .... \cap A_{r+1})$$

Hence if (4·6) is true for $n=r$, it is also true for $n = (r+1)$. But we have proved in (*) that (4·6) is true for $n=2$. Hence by the principle of mathematical induction, it follows that (4·6) is true for all positive integral values of $n$.

**Remarks. 1.** If we write

$$P(A_i) = p_i, P(A_i \cap A_j) = p_{ij}, P(A_i \cap A_j \cap A_k) = p_{ijk}$$

and so on and

$$S_1 = \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} P(A_i)$$

$$S_2 = \sum\sum_{1 \le i < j \le n} p_{ij} = \sum\sum_{1 \le i < j \le n} P(A_i \cap A_j)$$

$$S_3 = \sum\sum\sum_{1 \le i < j < k \le n} p_{ijk} \quad \text{and so on,}$$

then

$$P(\bigcup_{i=1}^{n} A_i) = S_1 - S_2 + S_3 - ... + (-1)^{n-1} S_n \quad ...(4\cdot 6a)$$

2. If all the events $A_i$, $(i = 1, 2, ..., n)$ are mutually disjoint then (4·6) gives

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

3. From practical point of view the theorem can be restated in a slightly different form. Let us suppose that an event $A$ can materialise in several mutually exclusive forms, *viz.*, $A_1, A_2, ..., A_n$ which may be regarded as that many mutually exclusive events. If $A$ happens then any one of the events $A_i$, $(i = 1, 2, ..., n)$ must happen and conversely if any one of the events $A_i$, $(i = 1, 2, ..., n)$ happens, then $A$ happens. Hence the probability of happening of $A$ is the same as the probability of happening of any one of its (unspecified) mutually exclusive forms. From this point of view, the total probability theorem can be restated as follows:

*The probability of happening of an event A is the sum of the probabilities of happening of its mutually exclusive forms $A_1, A_2, ..., A_n$. Symbolically,*

$$P(A) = P(A_1) + P(A_2) + ... + P(A_n) \quad (4\cdot 6b)$$

The probabilities $P(A_1), P(A_2), ..., P(A_n)$ of the mutually exclusive forms of $A$ are known as the *partial probabilities*. Since $P(A)$ is their sum, it may be called the *total probability* of $A$. Hence the name of the theorem.

**Theorem 4·7.** *(Boole's inequality). For n events $A_1, A_2, ..., A_n$, we have*

$$(a)\quad P\left(\bigcap_{i=1}^{n} A_i\right) \geq \sum_{i=1}^{n} P(A_i) - (n-1) \qquad ...(4\cdot7)$$

$$(b)\quad P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i) \qquad ...(4\cdot7a)$$

[Delhi Univ. B.Sc. (Stat Hons.), 1992, 1989]

**Proof.** *(a)* $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \leq 1$

$$\Rightarrow \quad P(A_1 \cap A_2) \geq P(A_1) + P(A_2) - 1 \qquad (*)$$

Hence (4·7) is true for $n = 2$.

Let us now suppose that (4·7) is true for $n = r$ (say), such that

$$P\left(\bigcap_{i=1}^{r} A_i\right) \geq \sum_{i=1}^{r} P(A_i) - (r-1) \qquad (**)$$

Then

$$P\left(\bigcap_{i=1}^{r+1} A_i\right) = P\left(\bigcap_{i=1}^{r} A_i \cap A_{r+1}\right)$$

$$\geq P\left(\bigcap_{i=1}^{r} A_i\right) + P(A_{r+1}) - 1 \qquad \text{[From } (*)]$$

$$\geq \sum_{i=1}^{r} P(A_i) - (r-1) + P(A_{r+1}) - 1 \qquad \text{[From } (**)]$$

$$\Rightarrow \quad P\left(\bigcap_{i=1}^{r+1} A_i\right) \geq \sum_{i=1}^{r+1} P(A_i) - r$$

$\Rightarrow$ (4·7) is true for $n = r + 1$ also.

The result now follows by the principle of mathematical induction.

(b) Applying the inequality (4·7) to the events $\overline{A}_1, \overline{A}_2, ..., \overline{A}_n$, we get

$$P(\overline{A}_1 \cap \overline{A}_2 \cap ... \cap \overline{A}_n) \geq [P(\overline{A}_1) + P(\overline{A}_2) + ... + P(\overline{A}_n)] - (n-1)$$

$$= [1 - P(A_1)] + [1 - P(A_2)] + ... + [1 - P(A_n)] - (n-1)$$

$$= 1 - P(A_1) - P(A_2) - ... - P(A_n)$$

$$\Rightarrow \quad P(A_1) + P(A_2) + .. + P(A_n) \geq 1 - P(\overline{A}_1 \cap \overline{A}_2 \cap ... \cap \overline{A}_n).$$

$$= 1 - P(\overline{A_1 \cup A_2 \cup ... \cup A_n}).$$

$$= P(A_1 \cup A_2 \cup ... \cup A_n)$$

$$\Rightarrow \quad P(A_1 \cup A_2 \cup ... \cup A_n) \leq P(A_1) + P(A_2) + ... + P(A_n)$$

as desired.

**Aliter for (b) i.e., (4·7a).** We have

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

$$\leq P(A_1) + P(A_2) \qquad [\because P(A_1 \cap A_2) \geq 0] \qquad \ldots(***)$$

Hence (4·7a) is true for $n = 2$.

Let us now suppose that (4·7a) is true for $n=r$, (say), so that

$$P\left(\bigcup_{i=1}^{r} A_i\right) \leq \sum_{i=1}^{r} P(A_i) \qquad \ldots(****)$$

Now

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) = P\left(\bigcup_{i=1}^{r} A_i \cup A_{r+1}\right)$$

$$\leq P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) \qquad \text{[Using (***)]}$$

$$\leq \sum_{i=1}^{r} P(A_i) + P(A_{r+1}) \qquad \text{[Using (****)]}$$

$$\Rightarrow \quad P\left(\bigcup_{i=1}^{r+1} A_i\right) \leq \sum_{i=1}^{r+1} P(A_i)$$

Hence if (4·7a) is true for $n=r$, then it is also true for $n=r+1$. But we have proved in (***) that (4·7a) is true for $n=2$. Hence by mathematical induction we conclude that (4·7a) is true for all positive integral values of $n$.

**Theorem 4·8.** *For n events* $A_1, A_2, \ldots, A_n$,

$$P\left[\bigcup_{i=1}^{n} A_i\right] \geq \sum_{i=1}^{n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j)$$

**[Delhi Univ. B.Sc. (Stat Hons.), 1986]**

**Proof.** We shall prove this theorem by the method of induction.

We know that

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3)$$
$$- [P(A_1 \cap A_2) + P(A_2 \cap A_3) + P(A_3 \cap A_1)] + P(A_1 \cap A_2 \cap A_3)$$

$$\Rightarrow P\left(\bigcup_{i=1}^{3} A_i\right) \geq \sum_{i=1}^{3} P(A_i) - \sum\sum_{1 \leq i < j \leq 3} P(A_i \cap A_j)$$

Thus the result is true for $n=3$. Let us now suppose that the result is true for $n=r$ (say), so that

$$P\left(\bigcup_{i=1}^{r} A_i\right) \geq \sum_{i=1}^{r} P(A_i) - \sum\sum_{1 \leq i < j \leq r} P(A_i \cap A_j) \qquad \ldots(*)$$

Now

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) = P\left(\bigcup_{i=1}^{r} A_i \cup A_{r+1}\right)$$

$$= P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) - P\left[\left(\bigcup_{i=1}^{r} A_i\right) \cap A_{r+1}\right]$$

$$= P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) - P\left[\bigcup_{i=1}^{r} (A_i \cap A_{r+1})\right]$$

$$\geq \left[\sum_{i=1}^{r} P(A_i) - \sum\sum_{1 \leq i < j < r} P(A_i \cap A_j)\right]$$

$$+ P(A_{r+1}) - P\left[\bigcup_{i=1}^{r} (A_i \cap A_{r+1})\right] \qquad \ldots(**)$$

[From (*)]

From Boole's inequality (*c.f.* Theorem 4·7 page 4·33), we get

$$P\left[\bigcup_{i=1}^{r} (A_i \cap A_{r+1})\right] \leq \sum_{i=1}^{r} P(A_i \cap A_{r+1})$$

$$\Rightarrow \quad -P\left[\bigcup_{i=1}^{r} (A_i \cap A_{r+1})\right] \geq -\sum_{i=1}^{r} P(A_i \cap A_{r+1})$$

$\therefore$ From (**), we get

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) \geq \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1 \leq i < j \leq r} P(A_i \cap A_j) - \sum_{i=1}^{r} P(A_i \cap A_{r+1})$$

$$\Rightarrow \quad P\left(\bigcup_{i=1}^{r+1} A_i\right) \geq \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1 \leq i < j \leq r+1} P(A_i \cap A_j)$$

Hence, if the theorem is true for $n = r$, it is also true for $n = r + 1$. But we have seen that the result is true for $n = 3$. Hence by mathematical induction, the result is true for all positive integral values of $n$.

## 4·7. Multiplication Law of Probability and Conditional Probability

**Theorem 4·8.** *For two events A and B*

$$\left.\begin{aligned} P(A \cap B) &= P(A) \, . \, P(B \mid A), \; P(A) > 0 \\ &= P(B) \, . \, P(A \mid B), \; P(B) > 0 \end{aligned}\right\} \qquad \ldots(4\cdot8)$$

*where $P(B \mid A)$ represents the conditional probability of occurrence of B when the event A has already happened and $P(A \mid B)$ is the conditional probability of happening of A, given that B has already happened.*

**Proof.**

$$P(A) = \frac{n(A)}{n(S)} \; ; \; P(B) = \frac{n(B)}{n(S)} \text{ and } P(A \cap B) = \frac{n(A \cap B)}{n(S)} \qquad (*)$$

For the conditional event $A \mid B$, the favourable outcomes must be one of the sample points of $B$, *i.e.*, for the event $A \mid B$, the sample space is $B$ and out of the $n(B)$ sample points, $n(A \cap B)$ pertain to the occurrence of the event $A$. Hence

$$P(A \mid B) = \frac{n(A \cap B)}{n(B)}$$

Rewriting (*), we get

$$P(A \cap B) = \frac{n(B)}{n(S)} \cdot \frac{n(A \cap B)}{n(B)} = P(B) \, . \, P(A \mid B)$$

Similarly we can prove :

$$P(A \cap B) = \frac{n(A)}{n(S)} \cdot \frac{n(A \cap B)}{n(A)} = P(A) \cdot P(B \mid A)$$

**Remarks. 1.** $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$ and $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

Thus the conditional probabilities $P(B \mid A)$ and $P(A \mid B)$ are defined if and only if $P(A) \neq 0$ and $P(B) \neq 0$, respectively.

2. *(i)* For $P(B) > 0$, $P(A \mid B) \leq P(A)$

   *(ii)* The conditional probability $P(A \mid B)$ is not defined if $P(B) = 0$.

   *(iii)* $P(B \mid B) = 1$.

**3. Multiplication Law of Probability for Independent Events.** If $A$ and $B$ are independent then

$$P(A \mid B) = P(A) \quad \text{and} \quad P(B \mid A) = P(B)$$

Hence (4·8) gives :

$$P(A \cap B) = P(A)\, P(B) \qquad \text{...(4·8}a)$$

provided $A$ and $B$ are independent.

**4·7·1. Extension of Multiplication Law of Probability.** *For $n$ events $A_1, A_2, \ldots, A_n$, we have*

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \ldots \times P(A_n \mid A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \qquad \text{...(4·8}b)$$

*where $P(A_i \mid A_j \cap A_k \cap \ldots \cap A_l)$ represents the conditional probability of the event $A_i$ given that the events $A_j, A_k, \ldots, A_l$ have already happened.*

**Proof.** We have for three events $A_1, A_2$, and $A_3$

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P[A_1 \cap (A_2 \cap A_3)] \\ &= P(A_1)\, P(A_2 \cap A_3 \mid A_1) \\ &= P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \end{aligned}$$

Thus we find that (4·8$b$) is true for $n=2$ and $n=3$. Let us suppose that (4·8$b$) is true for $n=k$, so that

$$P(A_1 \cap A_2 \cap \ldots \cap A_k) = P(A_1)\; P(A_2 \mid A_1)\; P(A_3 \mid A_1 \cap A_2) \ldots P(A_k \mid A_1 \cap A_2 \cap \ldots \cap A_{k-1})$$

Now

$$\begin{aligned} P[(A_1 \cap A_2 \cap \ldots \cap A_k) \cap A_{k+1}] &= P(A_1 \cap A_2 \cap \ldots \cap A_k) \\ &\quad \times P(A_{k+1} \mid A_1 \cap A_2 \cap \ldots \cap A_k) \\ &= P(A_1)\; P(A_2 \mid A_1) \ldots P(A_k \mid A_1 \cap A_2 \cap \ldots \cap A_{k-1}) \\ &\quad \times P(A_{k+1} \mid A_1 \cap A_2 \cap \ldots \cap A_k) \end{aligned}$$

Thus (4·8$b$) is true for $n=k+1$ also. Since (4·8$b$) is true for $n=2$ and $n=3$, by the principle of mathematical induction, it follows that (4·8$b$) is true for all positive integral values of $n$.

**Remark.** If $A_1, A_2, \ldots, A_n$ are independent events then

$$P(A_2 \mid A_1) = P(A_2),\; P(A_3 \mid A_1 \cap A_2) = P(A_3) \ldots P(A_n \mid A_1 \cap A_2 \cap \ldots \cap A_{n-1}) = P(A_n)$$

Hence (4·8*b*) gives :

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2) \dots \cap P(A_n), \qquad \dots(4\cdot8c)$$

provided $A_1, A_2, \dots, A_n$ are independent.

**Remark. Mutually Exclusive (Disjoint) Events and Independent Events.** Let $A$ and $B$ be mutually exclusive (disjoint) events with positive probabilities ($P(A) > 0, P(B) > 0$), *i.e.*, both $A$ and $B$ are possible events such that

$$A \cap B = \phi \quad \Rightarrow \quad P(A \cap B) = P(\phi) = 0 \qquad \dots(i)$$

Further, by compound probability theorem we have

$$P(A \cap B) = P(A) \,.\, P(B \mid A) = P(B) \,.\, P(A \mid B) \qquad \dots(ii)$$

Since $P(A) \neq 0$ ; $P(B) \neq 0$ , from *(i)* and *(ii)* we get

$$P(A \mid B) = 0 \neq P(A), \quad P(B \mid A) = 0 \neq P(B) \qquad \dots(iii)$$

$\Rightarrow$ $A$ and $B$ are dependent events.

Hence two possible mutually disjoint events are always dependent (not independent) events.

However, if $A$ and $B$ are independent events with $P(A) > 0$ and $P(B) > 0$, then

$$P(A \cap B) = P(A)\ P(B) \neq 0$$

$\Rightarrow$ $A$ and $B$ cannot be mutually exclusive.

*Hence two independent events (both of which are possible events), cannot be mutually disjoint.*

**4·7·2. Given n independent events $A_i$, (i =1,2,...,n) with respective probabilities of occurrence $p_i$, to find the probability of occurrence of at least one of them.**

We have

$$P(A_i) = p_i \quad \Rightarrow \quad P(\bar{A}_i) = 1 - p_i\,; \ i = 1, 2, \dots, n$$

$$[\because \overline{(A_1 \cup A_2 \cup \dots \cup A_n)} = (\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n) \ \ (\textit{De–Morgan's Law})]$$

Hence the probability of happening of at least one of the events is given by

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - P\overline{(A_1 \cup A_2 \cup \dots \cup A_n)} \qquad \dots(*)$$

$$= 1 - P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n)$$

$$= 1 - P(\bar{A}_1)\,P(\bar{A}_2) \dots P(\bar{A}_n) \qquad \dots(**)$$

[*c.f.* Theorem 4·14 page 4·41]

$$= 1 - [(1 - p_1)(1 - p_2) \dots (1 - p_n)]$$

$$= \left[ \sum_{i=1}^{n} p_i - \sum_{\substack{i,j=1 \\ i<j}}^{n}\sum (p_i p_j) + \sum_{\substack{i,j,k=1 \\ i<j<k}}^{n}\sum\sum (p_i p_j p_k) \dots + (-1)^{n-1} (p_1 p_2 \dots p_n) \right]$$

**Remark.** The results in (*) and (**) are very important and are used quite often in numerical problems. Result (*) stated in words gives:

$P$ [happening of at least one of the events $A_1, A_2, \dots, A_n$ ]

$= 1 - P$ (none of the events $A_1, A_2, \dots, A_n$ happens) or equivalently,

$P$ {none of the given events happens}

$= 1 - P$ {at least one of them happens}.

**Theorem 4·9.** *For any three events A, B and C*

$$P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$$

**Proof.** We have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\Rightarrow P[(A \cap C) \cup (B \cap C)] = P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)$$

Dividing both sides by $P(C)$, we get

$$\frac{P[(A \cap C) \cup (B \cap C)]}{P(C)} = \frac{P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)}{P(C)}, P(C) > 0$$

$$= \frac{P(A \cap C)}{P(C)} + \frac{P(B \cap C)}{P(C)} - \frac{P(A \cap B \cap C)}{P(C)}$$

$$\Rightarrow \quad \frac{P[(A \cup B) \cap C]}{P(C)} = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$$

$$\Rightarrow \quad P[(A \cup B) \mid C] = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$$

**Theorem 4·10.** *For any three events A, B and C*

$$P(A \cap \overline{B} \mid C) + P(A \cap B \mid C) = P(A \mid C)$$

**Proof.** $P(A \cap \overline{B} \mid C) + P(A \cap B \mid C)$

$$= \frac{P(A \cap \overline{B} \cap C)}{P(C)} + \frac{P(A \cap B \cap C)}{P(C)}$$

$$= \frac{P(A \cap \overline{B} \cap C) + P(A \cap B \cap C)}{P(C)}$$

$$= \frac{P(A \cap C)}{P(C)} = P(A \mid C)$$

**Theorem 4·11.** *For a fixed B with $P(B) > 0$, $P(A \mid B)$ is a probability function.* **[Delhi Univ. B.Sc. (Stat. Hons.), 1991; (Maths Hons.), 1992]**

**Proof.**

*(i)* $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)} \geq 0$

*(ii)* $P(S \mid B) = \dfrac{P(S \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$

*(iii)* If $\{A_n\}$ is any finite or infinite sequences of disjoint events, then

$$P[\bigcup_n A_n \mid B] = \frac{P[(\bigcup_n A_n) \cap B]}{P(B)} = \frac{P[(\bigcup_n A_n . B)]}{P(B)}$$

$$= \frac{\sum_n P(A_n B)}{P(B)} = \sum_n \left[\frac{P(A_n B)}{P(B)}\right] = \sum_n P(A_n \mid B)$$

Hence the theorem.

**Remark.** For given $B$ satisfying $P(B) > 0$, the conditional probability $P[\cdot|B]$ also enjoys the same properties as the unconditional probability.

For example, in the usual notations, we have:

(*i*) $P[\phi \mid B] = 0$

(*ii*) $P[\bar{A} \mid B] = 1 - P[A \mid B]$

(*iii*) $P[\bigcup_{i=1}^{n} A_i \mid B] = \sum_{i=1}^{n} P[A_i \mid B]$,

where $A_1, A_2, ..., A_n$ are mutually disjoint events.

(*iv*) $P(A_1 \cup A_2 \mid B) = P(A_1 \mid B) + P(A_2 \mid B) - P(A_1 A_2 \mid B)$

(*v*) If $E \subset F$, then $P(E \mid B) \leq P(F \mid B)$

and so on.

The proofs of results (*iv*) and (*v*) are given in theorems 4·9 and 4·13 respectively. Others are left as exercises to the reader.

**Theorem 4·12.** *For any three events, A, B and C defined on the sample space S such that $B \subset C$ and $P(A) > 0$,*

$$P(B \mid A) \leq P(C \mid A)$$

**Proof.**
$$P(C \mid A) = \frac{P(C \cap A)}{P(A)} \qquad \text{(By definition)}$$

$$= \frac{P[(B \cap C \cap A) \cup (\bar{B} \cap C \cap A)]}{P(A)}$$

$$= \frac{P[B \cap C \cap A)}{P(A)} + \frac{P(\bar{B} \cap C \cap A)}{P(A)} \qquad \text{(Using axiom 3)}$$

$$= P[(B \cap C|A) + (\bar{B} \cap C \cap A)]$$

Now
$$B \subset C \quad \Rightarrow \quad B \cap C = B$$

$$\therefore \qquad P(C \mid A) = P(B \mid A) + P(\bar{B} \cap C|A)$$

$$\Rightarrow \qquad P(C \mid A) \geq P(B \mid A)$$

**4·7·3. Independent Events.** *An event B is said to be independent (or statistically independent) of event A, if the conditional probability of B given A i.e., P (B | A) is equal to the unconditional probability of B, i.e., if*

$$P(B \mid A) = P(B)$$

Since

$$P(A \cap B) = P(B \mid A)\, P(A) = P(A \mid B)\, P(B)$$

and since $P(B \mid A) = P(B)$ when $B$ is independent of $A$, we must have $P(A \mid B) = P(A)$ or it follows that $A$ is also independent of $B$. Hence the events $A$ and $B$ are independent if and only if

$$P(A \cap B) = P(A)\, P(B) \qquad ...(4·9)$$

**4·7·4. Pairwise Independent Events**

**Definition.** *A set of events $A_1, A_2, ..., A_n$ are said to be pair-wise independent if*

$$P(A_i \cap A_j) = P(A_i)\, P(A_j) \ \forall \ i \neq j \qquad ...(4·10)$$

**4·7·5. Conditions for Mutual Independence of n Events.** Let $S$ denote the sample space for a number of events. The events in $S$ are said to be mutually indepdendent if the probability of the simultaneous occurrence of (any) finite number of them is equal to the product of their separate probabilities.

If $A_1, A_2, ..., A_n$ are $n$ events, then for their mutual independence, we should have

*(i)* $\quad P(A_i \cap A_j) = P(A_i) P(A_j), \; (i \neq j \; ; i, j = 1, 2, ..., n)$

*(ii)* $\quad P(A_i \cap A_j \cap A_k) = P(A_i) P(A_j) P(A_k), \; (i \neq j \neq k \; ; i ; j ; k = 1, 2, ..., n)$

$$\vdots \qquad\qquad \vdots$$

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1) P(A_2) ... P(A_n)$$

It is interesting to note that the above equations give respectively ${}^nC_2, {}^nC_3, ..., {}^nC_n$ conditions to be satisfied by $A_1, A_2, ..., A_n$.

Hence the total number of conditions for the mutual independence of $A_1, A_2, ..., A_n$ is ${}^nC_2 + {}^nC_3 + ... + {}^nC_n$.

Since ${}^nC_0 + {}^nC_1 + {}^nC_2 + ... + {}^nC_n = 2^n$, we get the required number of conditions as $(2^n - 1 - n)$.

In particular for three events $A_1, A_2$ and $A_3$, $(n = 3)$, we have the following $2^3 - 1 - 3 = 4$, conditions for their mutual independence.

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1) \; P(A_2) \\ P(A_2 \cap A_3) &= P(A_2) \; P(A_3) \\ P(A_1 \cap A_3) &= P(A_1) \; P(A_3) \\ P(A_1 \cap A_2 \cap A_3) &= P(A_1) \; P(A_2) \; P(A_3) \qquad ...(4·11) \end{aligned}$$

**Remarks. 1.** It may be observed that pairwise or mutual independence of events $A_1, A_2, ..., A_n$, is defined only when $P(A_i) \neq 0$, for $i = 1, 2, ..., n$.

**2.** If the events $A$ and $B$ are such that $P(A) \neq 0$, $P(B) \neq 0$ and $A$ is independent of $B$, then $B$ is independent of $A$.

**Proof.** We are given that

$$P(A \mid B) = P(A)$$

$$\Rightarrow \quad \frac{P(A \cap B)}{P(B)} = P(A)$$

$$\Rightarrow \quad P(A \cap B) = P(A) \; P(B)$$

$$\Rightarrow \quad \frac{P(B \cap A)}{P(A)} = P(B) \qquad [\because P(A) \neq 0 \text{ and } A \cap B = B \cap A]$$

$$\Rightarrow \quad P(B \mid A) = P(B),$$

which by definition of independent events, means that $B$ is independent of $A$.

**3.** It may be noted that pairwise independence of events does not imply their mutual independence. For illustrations, see Examples 4·50 and 4·51.

**Theorem 4·13.** *If $A$ and $B$ are independent events then $A$ and $\overline{B}$ are also independent events.*

**Proof.** By theorem 4·4, we have

$$P(A \cap \bar{B}) = P(A) - P(A \cap B)$$
$$= P(A) - P(A)P(B) \quad [\because A \text{ and } B \text{ are independent}]$$
$$= P(A)[1 - P(B)]$$
$$= P(A)\, P(\bar{B})$$

$\Rightarrow$ $A$ and $\bar{B}$ are independent events.

**Aliter.** $P(A \cap B) = P(A)\, P(B) = P(A)\, P(B \mid A) = P(B)\, P(A \mid B)$

*i.e.,* $P(B \mid A) = P(B) \Rightarrow B$ is independent of $A$.

also $P(A \mid B) = P(A) \Rightarrow A$ is independent of $B$.

Also $P(B \mid A) + P(\bar{B} \mid A) = 1 \Rightarrow P(B) + P(\bar{B} \mid A) = 1$

or $P(\bar{B} \mid A) = 1 - P(B) = P(\bar{B})$

$\therefore$ $\bar{B}$ is independent of $A$ and by symmetry we say that $A$ is independent of $\bar{B}$. Thus $A$ and $\bar{B}$ are independent events.

**Remark.** Similarly, we can prove that if $A$ and $B$ are independent events then $\bar{A}$ and $B$ are also independent events.

**Theorem 4·14.** *If $A$ and $B$ are independent events then $\bar{A}$ and $\bar{B}$ are also independent events.*

**Proof.** We are given $P(A \cap B) = P(A)\, P(B)$

Now $P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B)$

$$= 1 - [P(A) + P(B) - P(A \cap B)]$$
$$= 1 - [P(A) + P(B) - P(A)P(B)]$$
$$= 1 - P(A) - P(B) + P(A)P(B)$$
$$= [1 - P(B)] - P(A)[1 - P(B)]$$
$$= [1 - P(A)][1 - P(B)] = P(\bar{A})P(\bar{B})$$

$\therefore$ $\bar{A}$ and $\bar{B}$ are independent events.

**Aliter.** We know

$$P(\bar{A} \mid \bar{B}) + P(A \mid \bar{B}) = 1$$

$\Rightarrow$ $P(\bar{A} \mid \bar{B}) + P(A) = 1$ (*c.f.* Theorem 4·13)

$\Rightarrow$ $P(\bar{A} \mid \bar{B}) = 1 - P(A) = P(\bar{A})$

$\therefore$ $\bar{A}$ and $\bar{B}$ are independent events.

**Theorem 4·15.** *If $A, B, C$ are mutually independent events then $A \cup B$ and $C$ are also independent.*

**Proof.** We are required to prove:

$$P[(A \cup B) \cap C] = P(A \cup B)P(C)$$

L.H.S. $= P[(A \cap C) \cup (B \cap C)]$ [Distributive Law]

$$= P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)$$
$$= P(A)P(C) + P(B)P(C) - P(A)P(B)P(C)$$

[$\because A, B$ and $C$ are mutually independent]

$$= P(C)\,[P(A) + P(B) - P(A \cap B)]$$

$$= P(C)\ P(A \cup B) = \text{R.H.S.}$$

Hence $(A \cup B)$ and $C$ are independent.

**Theorem 4·16.** *If A, B and C are random events in a sample space and if A, B and C are pairwise independent and A is independent of ( B ∪ C ), then A, B and C are mutually independent.*

**Proof.** We are given

$$\left.\begin{aligned} P(A \cap B) &= P(A)P(B) \\ P(B \cap C) &= P(B)P(C) \\ P(A \cap C) &= P(A)P(C) \\ P[A \cap (B \cup C)] &= P(A)P(B \cup C) \end{aligned}\right\} \quad \ldots (*)$$

Now $\quad P[A \cap (B \cup C)] = P[(A \cap B) \cup (A \cap C)]$

$$= P(A \cap B) + P(A \cap C) - P[(A \cap B) \cap (A \cap C)]$$

$$= P(A).P(B) + P(A).P(C) - P(A \cap B \cap C) \quad \ldots (**)$$

and $P(A)P(B \cup C) = P(A)[P(B) + P(C) - P(B \cap C)]$

$$= P(A).P(B) + P(A)P(C) - P(A)P(B \cap C) \quad \ldots (***)$$

From (**) and (***), on using (*), we get

$$P(A \cap B \cap C) = P(A)P(B \cap C) = P(A)P(B)P(C)$$

Hence $A, B, C$ are mutually independent.

**Theorem 4·17.** *For any two events A and B,*

$$P(A \cap B) \le P(A) \le P(A \cup B) \le P(A) + P(B)$$

**[Patna Univ. B.A.(Stat. Hons.), 1992; Delhi Univ. B.Sc.(Stat. Hons.), 1989]**

**Proof.** We have

$$A = (A \cap \bar{B}) \cup (A \cap B)$$

Using axiom 3, we have

$$P(A) = P[(A \cap \bar{B}) \cup (A \cap B)] = P(A \cap \bar{B}) + P(A \cap B)$$

Now $\quad P[(A \cap \bar{B}) \ge 0 \quad$ (From axiom 1)

$\therefore \quad P(A) \ge P(A \cap B) \quad$ ..(*)

Similarly $\quad P(B) \ge P(A \cap B)$

$\Rightarrow \quad P(B) - P(A \cap B) \ge 0$

Now $\quad P(A \cup B) = P(A) + [P(B) - P(A \cap B)] \quad$ ...(**)

$\therefore \quad P(A \cup B) \ge P(A) \quad \Rightarrow \quad P(A) \le P(A \cup B) \quad$ ...(***)

Also $\quad P(A \cup B) \le P(A) + P(B) \quad$ [From (**)]

Hence from (*), (**) and (***), we get

$$P(A \cap B) \le P(A) \le P(A \cup B) \le P(A) + P(B)$$

**Aliter.** Since $A \cap B \subset A$, by Theorem 4·6 *(ii)* page 4·30, we get

$$P(A \cap B) \le P(A).$$

Also $A \subset (A \cup B) \quad \Rightarrow \quad P(A) \le P(A \cup B)$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\le P(A) + P(B) \qquad [\because P(A \cap B) \ge 0]$$

Combining the above results, we get

$$P(A \cap B) \le P(A) \le P(A \cup B) \le P(A) + P(B)$$

**Example 4·12.** *Two dice, one green and the other red, are thrown. Let A be the event that the sum of the points on the faces shown is odd, and B be the event of at least one ace (number '1').*

*(a) Describe the (i) complete sample space, (ii) events* $A$, $B$, $\overline{B}$, $A \cap B$, $A \cup B$, *and* $A \cap \overline{B}$ *and find their probabilities assuming that all the 36 sample points have equal probabilities.*

*(b) Find the probabilities of the events :*

*(i)* $(\overline{A} \cup \overline{B})$ *(ii)* $(\overline{A} \cap \overline{B})$ *(iii)* $(A \cap \overline{B})$ *(iv)* $(\overline{A} \cap B)$ *(v)* $(\overline{A \cap B})$ *(vi)* $(\overline{A} \cup B)$ *(vii)* $(\overline{A \cup B})$ *(viii)* $\overline{A} \cap (A \cup B)$ *(ix)* $A \cup (\overline{A} \cap B)$ *(x)* $(A \mid B)$ and $(B \mid A)$, and *(xi)* $(\overline{A} \mid \overline{B})$ *and* $(\overline{B} \mid \overline{A})$.

**Solution.** *(a)* The sample space consists of the 36 elementary events.

(1,1) ; (1,2) ; (1,3) ; (1,4) ; (1,5) ; (1,6)
(2,1) ; (2,2) ; (2,3) ; (2,4) ; (2,5) ; (2,6)
(3,1) ; (3,2) ; (3,3) ; (3,4) ; (3,5) ; (3,6)
(4,1) ; (4,2) ; (4,3) ; (4,4) ; (4,5) ; (4,6)
(5,1) ; (5,2) ; (5,3) ; (5,4) ; (5,5) ; (5,6)
(6,1) ; (6,2) ; (6,3) ; (6,4) ; (6,5) ; (6,6)

where, for example, the ordered pair (4, 5) refers to the elementary event that the green die shows 4 and and the red die shows 5.

$A$ = The event that the sum of the numbers shown by the two dice is odd.
= {(1,2); (2,1); (1,4); (2,3); (3,2); (4,1); (1,6); (2,5)
(3,4); (4,3); (5,2); (6,1); (3,6); (4,5); (5,4); (6,3)
(5,6); (6,5)} and therefore

$$P(A) = \frac{n(A)}{n(S)} = \frac{18}{36}$$

$B$ = The event that at least one face is 1,
= {(1,1) ; (1,2) ; (1,3) ; (1,4) ; (1,5) ; (1,6)
(2,1) ; (3,1) ; (4,1) ; (5,1) ; (6,1)} and therefore

$$P(B) = \frac{n(B)}{n(S)} = \frac{11}{36}$$

$\overline{B}$ = The event that each of the face obtained is not an ace.
= {(2,2); (2,3); (2,4); (2,5); (2,6); (3,2); (3,3);
(3,4) ; (3,5) ; (3,6); (4,2) ; (4,3); (4,4) ; (4,5);
(4,6) ; (5,2) ; (5,3); (5,4) ; (5,5); (5,6); (6,2) ;
(6,3) ; (6,4) ; (6,5); (6,6)} and therefore

$$P(\overline{B}) = \frac{n(\overline{B})}{n(S)} = \frac{25}{36}$$

$A \cap B$ = The event that sum is odd and at least one face is an ace.
= {(1,2); (2,1); (1,4); (4,1); (1,6); (6,1)}

$$\therefore \quad P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{6}{36} = \frac{1}{6}$$

$$A \cup B = \{(1,2);(2,1);(1,4);(2,3);(3,2);(4,1);(1,6);(2,5)\\ (3,4);(4,3);(5,2);(6,1);(3,6);(4,5);(5,4);(6,3)\\ (5,6);(6,5);(1,1);(1,3);(1,5);(3,1);(5,1)\}$$

$$\therefore \quad P(A \cup B) = \frac{n(A \cup B)}{n(S)} = \frac{23}{36}$$

$$A \cap \bar{B} = \{(2,3);(3,2);(2,5);(3,4);(3,6);(4,3);(4,5);(5,2)\\ (5,4);(5,6);(6,3);(6,5)\}$$

$$P(A \cap \bar{B}) = \frac{n(A \cap \bar{B})}{n(S)} = \frac{12}{36} = \frac{1}{3}$$

*(b) (i)* $\quad P(\bar{A} \cup \bar{B}) = P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - \frac{1}{6} = \frac{5}{6}$

*(ii)* $\quad P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - \frac{23}{36} = \frac{13}{36}$

*(iii)* $\quad P(A \cap \bar{B}) = P(A) - P(A \cap B) = \frac{18}{36} - \frac{6}{36} = \frac{12}{36} = \frac{1}{3}$

*(iv)* $\quad P(\bar{A} \cap B) = P(B) - P(A \cap B) = \frac{11}{36} - \frac{6}{36} = \frac{5}{36}$

*(v)* $\quad P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - \frac{1}{6} = \frac{5}{6}$

*(vi)* $\quad P(\bar{A} \cup B) = P(\bar{A}) + P(B) - P(\bar{A} \cap B)$

$$= \left(1 - \frac{18}{36}\right) + \frac{11}{36} - \frac{5}{36} = \frac{2}{3}$$

*(vii)* $\quad P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - \frac{23}{36} = \frac{13}{36}$

*(viii)* $\quad P[\bar{A} \cap (A \cup B)] = P[(A \cap \bar{A}) \cup (\bar{A} \cap B)]$

$$= P(\bar{A} \cap B) = \frac{5}{36} \qquad [\because A \cap \bar{A} = \phi]$$

*(ix)* $\quad P[A \cup (\bar{A} \cap B)] = P(A) + P(\bar{A} \cap B) - P(A \cap \bar{A} \cap B)$

$$= P(A) + P(\bar{A} \cap B) = \frac{18}{36} + \frac{5}{36} = \frac{23}{36}$$

*(x)* $\quad P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{6/36}{11/36} = \frac{6}{11}$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{6/36}{18/36} = \frac{6}{18} = \frac{1}{3}$$

*(xi)* $\quad P(\bar{A} \mid \bar{B}) = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{B})} = \frac{13/36}{25/36} = \frac{13}{25}$

$$P(\bar{B} \mid \bar{A}) = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{A})} = \frac{13/36}{18/36} = \frac{13}{18}$$

**Example 4·13.** *If two dice are thrown, what is the probability that the sum is (a) greater than 8, and (b) neither 7 nor 11?*

**Solution.** *(a)* If $S$ denotes the sum on the two dice, then we want $P(S > 8)$. The required event can happen in the following mutually exclusive ways:

*(i)* $S = 9$ *(ii)* $S = 10$ *(iii)* $S = 11$ *(iv)* $S = 12$.

Hence by addition theorem of probability

$$P(S > 8) = P(S = 9) + P(S = 10) + P(S = 11) + P(S = 12)$$

In a throw of two dice, the sample space contains $6^2 = 36$ points.

The number of favourable cases can be enumerated as follows:

$S = 9$ : (3, 6), (6, 3), (4, 5), (5, 4), *i.e.*, 4 sample points.

$$\therefore \quad P(S=9) = \frac{4}{36}$$

$S = 10$ : (4,6), (6,4), (5,5), *i.e.*, 3 sample points.

$$\therefore \quad P(S=10) = \frac{3}{36}$$

$S = 11$ : (5, 6), (6, 5), *i.e.*, 2 sample points.

$$\therefore \quad P(S=11) = \frac{2}{36}$$

$S = 12$ : (6, 6), *i.e.*, 1 sample point.

$$\therefore \quad P(S=12) = \frac{1}{36}$$

$$\therefore \quad P(S > 8) = \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{10}{36} = \frac{5}{18}$$

*(b)* Let $A$ denote the event of getting the sum of 7 and $B$ denote the event of getting the sum of 11 with a pair of dice.

$S = 7$ : (1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3), *i.e.*, 6 distinct sample points.

$$\therefore \quad P(A) = P(S=7) = \frac{6}{36} = \frac{1}{6}$$

$S = 11$ : (5, 6), (6, 5), $P(B) = P(S=11) = \dfrac{2}{36} = \dfrac{1}{18}$

$$\therefore \text{ Required probability} = P(\bar{A} \cap \bar{B}) = 1 - P(A \cup B)$$

$$= 1 - [P(A) + P(B)]$$

( $\because$ $A$ and $B$ are disjoint events )

$$= 1 - \frac{1}{6} - \frac{1}{18} = \frac{7}{9}$$

**Example 4·14.** *An urn contains 4 tickets numbered 1, 2, 3, 4 and another contains 6 tickets numbered 2, 4, 6, 7, 8, 9. If one of the two urns is chosen at random and a ticket is drawn at random from the chosen urn, find the probabilities that the ticket drawn bears the number (i) 2 or 4, (ii) 3, (iii) 1 or 9*

**[Calicut Univ. B.Sc., 1992]**

**Solution.** *(i)* Required event can happen in the following mutually exclusive ways:

*(I)* First urn is chosen and then a ticket is drawn.

*(II)* Second urn is chosen and then a ticket is drawn.

Since the probability of choosing any urn is ½, the required probability '$p$' is given by

$$p = P(I) + P(II)$$

$$= \frac{1}{2} \times \frac{2}{4} + \frac{1}{2} \times \frac{2}{6} = \frac{5}{12}$$

*(ii)* Required probability $= \frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times 0 = \frac{1}{8}$

($\because$ in the 2nd urn there is no ticket with number 3)

*(iii)* Required probability $= \frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{6} = \frac{5}{24}$

**Example 4·15.** *A card is drawn from a well-shuffled pack of playing cards. What is the probability that it is either a spade or an ace ?*

**Solution.** The equiprobable sample space $S$ of drawing a card from a well-shuffled pack of playing cards consists of 52 sample points.

If $A$ and $B$ denote the events of drawing a 'spade card' and 'an ace' respectively then $A$ consists of 13 sample points and $B$ consists of 4 sample points so that,

$$P(A) = \frac{13}{52} \quad \text{and} \quad P(B) = \frac{4}{52}$$

The compound event $A \cap B$ consists of only one sample point, *viz.*, ace of spade so that,

$$P(A \cap B) = \frac{1}{52}$$

The probability that the card drawn is either a spade or an ace is given by

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{4}{13}$$

**Example 4·16.** *A box contains 6 red, 4 white and 5 black balls. A person draws 4 balls from the box at random. Find the probability that among the balls drawn there is at least one ball of each colour.* **(Nagpur Univ. B.Sc., 1992)**

**Solution.** The required event $E$ that 'in a draw of 4 balls from the box at random there is at least one ball of each colour', can materialise in the following mutually disjoint ways :

*(i)* 1 Red, 1 White, 2 Black balls

*(ii)* 2 Red, 1 White, 1 Black balls

*(iii)* 1 Red, 2 White, 1 Black balls.

Hence by the addition theorem of probability, the required probability is given by

$$P(E) = P(i) + P(ii) + P(iii)$$
$$= \frac{{}^6C_1 \times {}^4C_1 \times {}^5C_2}{{}^{15}C_4} + \frac{{}^6C_2 \times {}^4C_1 \times {}^5C_1}{{}^{15}C_4} + \frac{{}^6C_1 \times {}^4C_2 \times {}^5C_1}{{}^{15}C_4}$$
$$= \frac{1}{{}^{15}C_4}\left[6 \times 4 \times 10 + 15 \times 4 \times 5 + 6 \times 6 \times 5\right]$$
$$= \frac{4!}{15 \times 14 \times 13 \times 12}\left[240 + 300 + 180\right]$$
$$= \frac{24 \times 720}{15 \times 14 \times 13 \times 12} = 0{\cdot}5275$$

**Example 4·17.** *Why does it pay to bet consistently on seeing 6 at least once in 4 throws of a die, but not on seeing a double six at least once in 24 throws with two dice? (de Mere's Problem).*

**Solution.** The probability of getting a '6' in a throw of die $= 1/6$.

$\therefore$ The probability of not getting a '6' in a throw of die

$$= 1 - 1/6 = 5/6.$$

By compound probability theorem, the probability that in 4 throws of a die no '6' is obtained $= (5/6)^4$

Hence the probability of obtaining '6' at least once in 4 throws of a die $= 1 - (5/6)^4 = 0{\cdot}516$

Now, if a trial consists of throwing two dice at a time, then the probability of getting a 'double' of '6' in a trial $= 1/36$.

Thus the probability of not getting a 'double of 6' in a trial $= 35/36$.

The probability that in 24 throws, with two dice each, no 'double of 6' is obtained $= (35/36)^{24}$

Hence the probability of getting a 'double of 6' at least once in 24 throws $= 1 - (35/36)^{24} = 0{\cdot}491$.

Since the probability in the first case is greater than the probability in the second case, the result follows.

**Example 4·18.** *A problem in Statistics is given to the three students A, B and C whose chances of solving it are* $1/2$, $3/4$, *and* $1/4$ *respectively.*

*What is the probability that the problem will be solved if all of them try independently?* **[Madurai Kamraj Univ. B.Sc.,1986; Delhi Univ. B.A.,1991]**

**Solution.** Let $A$, $B$, $C$ denote the events that the problem is solved by the students $A, B, C$ respectively. Then

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{3}{4} \quad \text{and} \quad P(C) = \frac{1}{4}$$

The problem will be solved if at least one of them solves the problem. Thus we have to calculate the probability of occurrence of at least one of the three events $A, B, C$, *i.e.*, $P(A \cup B \cup C)$.

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ &\qquad - P(B \cap C) + P(A \cap B \cap C) \\ &= P(A) + P(B) + P(C) - P(A)\,P(B) - P(A)\,P(C) \\ &\qquad - P(B)\,P(C) + P(A)\,P(B)\,P(C) \end{aligned}$$

$(\because A, B, C$ are independent events.$)$

$$\begin{aligned} &= \frac{1}{2} + \frac{3}{4} + \frac{1}{4} - \frac{1}{2}\cdot\frac{3}{4} - \frac{3}{4}\cdot\frac{1}{4} \\ &\qquad - \frac{1}{2}\cdot\frac{1}{4} + \frac{1}{2}\cdot\frac{3}{4}\cdot\frac{1}{4} \\ &= \frac{29}{32} \end{aligned}$$

**Aliter.** $P(A \cup B \cup C) = 1 - P(\overline{A \cup B \cup C})$

$$= 1 - P(\overline{A} \cap \overline{B} \cap \overline{C})$$
$$= 1 - P(\overline{A})P(\overline{B})P(\overline{C})$$
$$= 1 - \left(1 - \frac{1}{2}\right)\left(1 - \frac{3}{4}\right)\left(1 - \frac{1}{4}\right)$$
$$= \frac{29}{32}$$

**Example 4·19.** *If* $A \cap B = \phi$, *then show that* ...(*)

$$P(A) \leq P(\overline{B})$$

**[Delhi Univ. B.Sc. (Maths Hons.) 1987]**

**Solution .** We have

$$A = (A \cap B) \cup (A \cap \overline{B})$$
$$= \phi \cup (A \cap \overline{B}) \quad \text{[Using * ]}$$
$$= A \cap \overline{B}$$
$$\Rightarrow \quad A \subseteq \overline{B}$$
$$\Rightarrow \quad P(A) \leq P(\overline{B})$$

as desired.

**Aliter.** Since $A \cap B = \phi$, we have $A \subset \overline{B}$, which implies that $P(A) \leq P(\overline{B})$.

**Example 4·20.** *Let A and B be two events such that*

$$P(A) = \frac{3}{4} \text{ and } P(B) = \frac{5}{8}$$

*show that*

*(a)* $P(A \cup B) \geq \dfrac{3}{4}$

*(b)* $\dfrac{3}{8} \leq P(A \cap B) \leq \dfrac{5}{8}$

**[Delhi Univ. B.Sc. Stat (Hons.) 1986,1988]**

**Solution.** *(i)* We have

$$A \subset (A \cup B)$$
$$\Rightarrow \quad P(A) \leq P(A \cup B)$$
$$\Rightarrow \quad \frac{3}{4} \leq P(A \cup B)$$
$$\Rightarrow \quad P(A \cup B) \geq \frac{3}{4}$$

*(ii)* $A \cap B \subseteq B$

$$\Rightarrow \quad P(A \cap B) \leq P(B) = \frac{5}{8} \quad ...(i)$$

Also $\quad P(A \cup B) = P(A) + P(B) - P(A \cap B) \leq 1$

$$\Rightarrow \quad \frac{3}{4} + \frac{5}{8} - 1 \leq P(A \cap B)$$

$$\Rightarrow \qquad \frac{6+5-8}{8} \le P(A \cap B)$$

$$\Rightarrow \qquad \frac{3}{8} \le P(A \cap B) \qquad \qquad ...(ii)$$

From *(i)* and *(ii)* we get

$$\frac{3}{8} \le P(A \cap B) \le \frac{5}{8}$$

**Example 4·21.** *(Chebychev's Problem). What is the chance that two numbers, chosen at random, will be prime to each other ?*

**Solution.** If any number '$a$' is divided by a prime number '$r$', then the possible remainders are 0, 1, 2, ...$r-1$. Hence the chance that '$a$' is divisible by $r$ is $1/r$ (because the only case favourable to this is remainder being 0). Similarly, the probability that any number '$b$' chosen at random is divisible by $r$ is $1/r$. Since the numbers $a$ and $b$ are chosen at random, the probability that none of them is divisible by '$r$' is given (by compound probability theorem) by :

$$\left(1-\frac{1}{r}\right) \times \left(1-\frac{1}{r}\right) = \left(1-\frac{1}{r}\right)^2; \quad r = 2, 3, 5, 7, \ldots$$

Hence the required probability that the two numbers chosen at random are prime to each other is given by

$$P = \prod_r \left(1-\frac{1}{r}\right)^2, \quad \text{where } r \text{ is a prime number.}$$

$$= \frac{6}{\pi^2} \qquad \qquad \text{(From trigonometry)}$$

**Example 4·22.** *A bag contains 10 gold and 8 silver coins. Two successive drawings of 4 coins are made such that : (i) coins are replaced before the second trial, (ii) the coins are not replaced before the second trial. Find the probability that the first drawing will give 4 gold and the second 4 silver coins.*

**[Allahabad Univ. B.Sc., 1987]**

**Solution.** Let $A$ denote the event of drawing 4 gold coins in the first draw and $B$ denote the event of drawing 4 silver coins in the second draw. Then we have to find the probability of $P(A \cap B)$.

*(i) Draws with replacement.* If the coins drawn in the first draw are replaced back in the bag before the second draw then the events $A$ and $B$ are independent and the required probability is given (using the multiplication rule of probability) by the expression

$$P(A \cap B) = P(A) \cdot P(B) \qquad \qquad ...(*)$$

*Ist draw.* Four coins can be drawn out of 10+8=18 coins in $^{18}C_4$ ways, which gives the exhaustive number of cases. In order that all these coins are of gold, they must be drawn out of the 10 gold coins and this can be done in $^{10}C_4$ ways. Hence

$$P(A) = {}^{10}C_4 \,/\, {}^{18}C_4$$

*2nd draw.* When the coins drawn in the first draw are replaced before the 2nd draw, the bag contains 18 coins. The probability of drawing 4 silver coins in the 2nd draw is given by $P(B) = {}^8C_4 \,/\, {}^{18}C_4$.

Substituting in (*), we have

$$P(A \cap B) = \frac{{}^{10}C_4}{{}^{18}C_4} \times \frac{{}^8C_4}{{}^{18}C_4}$$

*(ii) Draws without replacement.* If the coins drawn are not eplaced back before the second draw, then the events $A$ and $B$ are not independent and the required probability is given by

$$P(A \cap B) = P(A) \,.\, P(B \mid A) \qquad \ldots(**)$$

As discussed in part *(i)*, $P(A) = {}^{10}C_4 \,/\, {}^{18}C_4$.

Now, if the 4 gold coins which were drawn in the first draw are not replaced back, there are $18 - 4 = 14$ coins left in the bag and $P(B \mid A)$ is the probability of drawing 4 silver coins from the bag containing 14 coins out of which 6 are gold coins and 8 are silver coins.

Hence $\quad P(B \mid A) = {}^8C_4 \,/\, {}^{14}C_4$

Substituting in (**) we get

$$P(A \cap B) = \frac{{}^{10}C_4}{{}^{18}C_4} \times \frac{{}^8C_4}{{}^{14}C_4}$$

**Example 4·23.** *A consignment of 15 record players contains 4 defectives. The record players are selected at random, one by one, and examined. Those examined are not put back. What is the probability that the 9th one examined is the last defective?*

**Solution.** Let $A$ be the event of getting exactly 3 defectives in examination of 8 record players and let $B$ the event that the 9th piece examined is a defective one.

Since it is a problem of sampling without replacement and since there are 4 defectives out of 15 record players, we have

$$P(A) = \frac{\binom{4}{3} \times \binom{11}{5}}{\binom{15}{8}}$$

$P(B \mid A)$ = Probability that the 9th examined record player is defective given that there were 3 defectives in the first 8 pieces examined.

$= 1/7,$

since there is only one defective piece left among the remaining $15 - 8 = 7$ record players.

Hence the required probability is

$$P(A \cap B) = P(A) \,.\, P(B \mid A)$$

$$= \frac{\binom{4}{3} \times \binom{11}{5}}{\binom{15}{8}} \times \frac{1}{7} = \frac{8}{195}$$

**Example 4·24.** *p is the probability that a man aged x years will die in a year. Find the probability that out of n men $A_1, A_2, ..., A_n$ each aged x, $A_1$ will die in a year and will be the first to die.* **[Delhi Univ. B.Sc., 1985]**

**Solution.** Let $E_i$, $(i = 1, 2, ..., n)$ denote the event that $A_i$ dies in a year. Then

$$P(E_i) = p, (i = 1, 2, ..., n) \quad \text{and} \quad P(\overline{E}_i) = 1 - p.$$

The probability that none of $n$ men $A_1, A_2, ..., A_n$ dies in a year

$$= P(\overline{E}_1 \cap \overline{E}_2 \cap \ldots \cap \overline{E}_n) = P(\overline{E}_1) P(\overline{E}_2) \ldots P(\overline{E}_n)$$

(By compound probability theorem)

$$= (1 - p)^n$$

$\therefore$ The probability that at least one of $A_1, A_2, ..., A_n$, dies in a year

$$= 1 - P(\overline{E}_1 \cap \overline{E}_2 \cap \ldots \cap \overline{E}_n) = 1 - (1 - p)^n$$

The probability that among $n$ men, $A_1$ is the first to die is $1/n$ and since this event is independent of the event that at least one man dies in a year, required probability is

$$\frac{1}{n}\left[1 - (1 - p)^n\right]$$

**Example 4·25.** *The odds against Manager X settling the wage dispute with the workers are 8:6 and odds in favour of manager Y settling the same dispute are 14:16.*

*(i) What is the chance that neither settles the dispute, if they both try, independently of each other?*

*(ii) What is the probability that the dispute will be settled?*

**Solution.** Let $A$ be the event that the manager $X$ will settle the dispute and $B$ be the event that the Manager $Y$ will settle the dispute. Then clearly

$$P(\overline{A}) = \frac{8}{8+6} = \frac{4}{7} \quad \Rightarrow \quad P(A) = 1 - P(\overline{A}) = \frac{6}{14} = \frac{3}{7}$$

$$P(B) = \frac{14}{14+16} = \frac{7}{15} \quad \Rightarrow \quad P(\overline{B}) = 1 - P(B) = \frac{16}{14+16} = \frac{8}{15}$$

The required probability that neither settles the dispute is given by :

$$P(\overline{A} \cap \overline{B}) = P(\overline{A}) \times P(\overline{B}) = \frac{4}{7} \times \frac{8}{15} = \frac{32}{105}$$

[Since $A$ and $B$ are independent $\Rightarrow$ $\overline{A}$ and $\overline{B}$ are also independent]

*(ii)* The dispute will be settled if at least one of the managers $X$ and $Y$ settles the dispute. Hence the required probability is given by:

$$P(A \cup B) = \text{Prob. [ At least one of X and Y settles the dispute]}$$

$$= 1 - \text{Prob. [ None settles the dispute]}$$

$$= 1 - P(\overline{A} \cap \overline{B}) = 1 - \frac{32}{105} = \frac{73}{105}$$

**Example 4·26.** *The odds that person X speaks the truth are 3:2 and the odds that person Y speaks the truth are 5:3. In what percentage of cases are they likely to contradict each other on an identical point.*

**Solution.** Let us define the events:

$A$ : $X$ speaks the truth, $\quad B$ : $Y$ speaks the truth

Then $\bar{A}$ and $\bar{B}$ represent the complementary events that $X$ and $Y$ tell a lie respectively. We are given:

$$P(A) = \frac{3}{3+2} = \frac{3}{5} \quad \Rightarrow \quad P(\bar{A}) = 1 - \frac{3}{5} = \frac{2}{5}$$

and $$P(B) = \frac{5}{5+3} = \frac{5}{8} \quad \Rightarrow \quad P(\bar{B}) = 1 - \frac{5}{8} = \frac{3}{8}$$

The event $E$ that $X$ and $Y$ contradict each other on an identical point can happen in the following mutually exclusive ways:

*(i)* $X$ speaks the truth and $Y$ tells a lie, *i.e.*, the event $A \cap \bar{B}$ happens,

*(ii)* $X$ tells a lie and $Y$ speaks the truth, *i.e.*, the event $\bar{A} \cap B$ happens.

Hence by addition theorem of probability the required probability is given by:

$$P(E) = P(i) + P(ii) = P(A \cap \bar{B}) + P(\bar{A} \cap B)$$

$$= P(A) \, . \, P(\bar{B}) + P(\bar{A}) \, . P(B),$$

[Since $A$ and $B$ are independent ]

$$= \frac{3}{5} \times \frac{3}{8} + \frac{2}{5} \times \frac{5}{8} = \frac{19}{40} = 0{\cdot}475$$

Hence $A$ and $B$ are likely to contradict each other on an identical point in 47·5% of the cases.

**Example 4·27.** *A special dice is prepared such that the probabilities of throwing 1, 2, 3, 4, 5 and 6 points are :*

$$\frac{1-k}{6}, \frac{1+2k}{6}, \frac{1-k}{6}; \quad \frac{1+k}{6}, \frac{1-2k}{6}, \text{ and } \frac{1+k}{6}$$

*respectively. If two such dice are thrown, find the probability of getting a sum equal to 9.* **[Delhi Univ. B.Sc. (Stat. Hons.), 1988]**

**Solution.** Let $(x, y)$ denote the numbers obtained in a thrown of two dice, $x$ denoting the number on the first dice and $y$ denoting the number on the second dice. The sum $S = x+y = 9$, can be obtained in the following mutually disjoint ways:

*(i)* (3, 6), *(ii)* (6, 3), *(iii)* (4, 5), *(iv)* (5, 4)

Hence by addition theorem of probability:

$$P(S=9) = P(3,6) + P(6,3) + P(4,5) + P(5,4)$$

$$= P(x=3)\, P(y=6) + P(x=6)\, P(y=3) + P(x=4)\, P(y=5) + P(x=5)\, P(y=4),$$

since the number on one dice is independent of the number on the other dice.

$$\therefore \quad P(S=9)=\frac{(1-k)}{6}\cdot\frac{(1+k)}{6}+\frac{(1+k)}{6}\cdot\frac{(1-k)}{6}+\frac{(1+k)}{6}\cdot\frac{(1-2k)}{6}+\frac{(1-2k)}{6}\cdot\frac{(1+k)}{6}$$

$$=2\left(\frac{1+k}{36}\right)[(1-k)+(1-2k)]$$

$$=\frac{1}{18}(1+k)(2-3k)$$

**Example 4·28.** *(a) A and B alternately cut a pack of cards and the pack is shuffled after each cut. If A starts and the game is continued until one cuts a diamond, what are the respective chances of A and B first cutting a diamond?*

*(b) One shot is fired from each of the three guns. $E_1$, $E_2$, $E_3$ denote the events that the target is hit by the first, second and third gun respectively. If $P(E_1)=0{\cdot}5$, $P(E_2)=0{\cdot}6$ and $P(E_3)=0{\cdot}8$ and, $E_1, E_2, E_3$ are independent events, find the probability that (a) exactly one hit is registered, (b) at least two hits are registered.*

**Solution.** *(a)* Let $E_1$ and $E_2$, denote the events of $A$ and $B$ cutting a diamond respectively. Then

$$P(E_1)=P(E_2)=\frac{13}{52}=\frac{1}{4} \quad\Rightarrow\quad P(\overline{E_1})=P(\overline{E_2})=\frac{3}{4}$$

If $A$ starts the game, he can first cut the diamond in the following mutually exclusive ways:

*(i)* $E_1$ happens, *(ii)* $\overline{E_1}\cap\overline{E_2}\cap E_1$ happens, *(iii)* $\overline{E_1}\cap\overline{E_2}\cap\overline{E_1}\cap\overline{E_2}\cap E_1$ happens, and so on. Hence by addition theorem of probability, the probability '$p$' that $A$ first wins is given by

$$p = P(i)+P(ii)+P(iii)+\ldots\ldots$$

$$=P(E_1)+P(\overline{E_1}\cap\overline{E_2}\cap E_1)+P(\overline{E_1}\cap\overline{E_2}\cap\overline{E_1}\cap\overline{E_2}\cap E_1)+\ldots$$

$$=P(E_1)+P(\overline{E_1})P(\overline{E_2})P(E_1)+P(\overline{E_1})\cdot P(\overline{E_2})P(\overline{E_1})P(\overline{E_2})P(E_1)+\ldots$$

*(By Compound Probability Theorem)*

$$=\frac{1}{4}+\frac{3}{4}\cdot\frac{3}{4}\cdot\frac{1}{4}+\frac{3}{4}\cdot\frac{3}{4}\cdot\frac{3}{4}\cdot\frac{3}{4}\cdot\frac{1}{4}+\ldots\ldots$$

$$=\frac{\frac{1}{4}}{1-\frac{9}{16}}=\frac{4}{7}$$

The probability that $B$ first cuts a diamond

$$=1-p=1-\frac{4}{7}=\frac{3}{7}$$

*(b)* We are given

$$P(\overline{E_1})=0{\cdot}5,\ P(\overline{E_2})=0{\cdot}4 \text{ and } P(\overline{E_3})=0{\cdot}2$$

*(a)* Exactly one hit can be registered in the following mutually exclusive ways: *(i)* $E_1 \cap \overline{E_2} \cap \overline{E_3}$ happens, *(ii)* $\overline{E_1} \cap E_2 \cap \overline{E_3}$ happens, *(iii)* $\overline{E_1} \cap \overline{E_2} \cap E_3$ happens.

Hence by addition probability theorem, the required probability '$p$' is given by :

$$p = P(E_1 \cap \overline{E_2} \cap \overline{E_3}) + P(\overline{E_1} \cap E_2 \cap \overline{E_3}) + P(\overline{E_1} \cap \overline{E_2} \cap E_3)$$

$$= P(E_1)\, P(\overline{E_2})\, P(\overline{E_3}) + P(\overline{E_1})\, P(E_2)\, P(\overline{E_3}) + P(\overline{E_1})\, P(\overline{E_2})\, P(E_3)$$

(Since $E_1, E_2$ and $E_3$ are independent)

$$= 0{\cdot}5 \times 0{\cdot}4 \times 0{\cdot}2 + 0{\cdot}5 \times 0{\cdot}6 \times 0{\cdot}2 + 0{\cdot}5 \times 0{\cdot}4 \times 0{\cdot}8 = 0{\cdot}26.$$

*(b)* At least two hits can be registered in the following mutually exclusive ways:

*(i)* $E_1 \cap E_2 \cap \overline{E_3}$ happens *(ii)* $E_1 \cap \overline{E_2} \cap E_3$ happens, *(iii)* $\overline{E_1} \cap E_2 \cap E_3$ happens. *(iv)* $E_1 \cap E_2 \cap E_3$ happens.

Required probability

$$= P(E_1 \cap E_2 \cap \overline{E_3}) + P(E_1 \cap \overline{E_2} \cap E_3) + P(\overline{E_1} \cap E_2 \cap E_3) + P(E_1 \cap E_2 \cap E_3)$$

$$= 0{\cdot}5 \times 0{\cdot}6 \times 0{\cdot}2 + 0{\cdot}5 \times 0{\cdot}4 \times 0{\cdot}8 + 0{\cdot}5 \times 0{\cdot}6 \times 0{\cdot}8 + 0{\cdot}5 \times 0{\cdot}6 \times 0{\cdot}8$$

$$= 0{\cdot}06 + 0{\cdot}16 + 0{\cdot}24 + 0{\cdot}24 = 0{\cdot}70$$

**Example 4·29.** *Three groups of children contain respectively 3 girls and 1 boy, 2 girls and 2 boys, and 1 girl and 3 boys. One child is selected at random from each group. Show that the chance that the three selected consist of 1 girl and 2 boys is 13/32.* **[Madurai Univ. B.Sc.,1988; Nagpur Univ. B.Sc.,1991]**

**Solution.** The required event of getting 1 girl and 2 boys among the three selected children can materialise in the following three mutually disjoint cases:

| Group No. → | I | II | III |
|---|---|---|---|
| *(i)* | Girl | Boy | Boy |
| *(ii)* | Boy | Girl | Boy |
| *(iii)* | Boy | Boy | Girl |

Hence by addition theorem of probability,

Required probability $= P\,(i) + P\,(ii) + P\,(iii)$ ...(*)

Since the probability of selecting a girl from the first group is 3/4, of selecting a boy from the second is 2/4, and of selecting a boy from the third group is 3/4, and since these three events of selecting children from three groups are independent of each other, by compound probability theorem, we have

$$P\,(i) = \frac{3}{4} \times \frac{2}{4} \times \frac{3}{4} = \frac{9}{32}$$

Similarly, we have

$$P\,(ii) = \frac{1}{4} \times \frac{2}{4} \times \frac{3}{4} = \frac{3}{32}$$

and $$P\,(iii) = \frac{1}{4} \times \frac{2}{4} \times \frac{1}{4} = \frac{1}{32}$$

Substituting in ( *), we get

Required probability = $\frac{9}{32} + \frac{3}{32} + \frac{1}{32} = \frac{13}{32}$

## EXERCISE 4 (b)

1. *(a)* Which function defines a probability space on $S = (e_1, e_2, e_3)$

*(i)* $P(e_1) = \frac{1}{4}$, $P(e_2) = \frac{1}{3}$, $P(e_3) = \frac{1}{2}$

*(ii)* $P(e_1) = \frac{2}{3}$, $P(e_2) = -\frac{1}{3}$, $P(e_3) = \frac{2}{3}$

*(iii)* $P(e_1) = \frac{1}{4}$, $P(e_2) = \frac{1}{3}$, $P(e_3) = \frac{1}{2}$, and

*(iv)* $P(e_1) = 0$, $P(e_2) = \frac{1}{3}$, $P(e_3) = \frac{2}{3}$

**Ans.** *(i)* No, *(ii)* No, *(iii)* No, and *(iv)* Yes

*(b)* Let $S = (e_1, e_2, e_3, e_4)$, and let $P$ be a probability function on $S$.

*(i)* Find $P(e_1)$, if $P(e_2) = \frac{1}{3}, P(e_3) = \frac{1}{6}, P(e_4) = \frac{1}{9}$,

*(ii)* Find $P(e_1)$ and $P(e_2)$ if $P(e_3) = P(e_4) = \frac{1}{4}$ and $P(e_1) = 2P(e_2)$, and

*(iii)* Find $P(e_1)$ if $P[(e_2, e_3)] = \frac{2}{3}, P[(e_2, e_4)] = \frac{1}{2}$ and $P(e_2) = \frac{1}{3}$.

**Ans.** *(i)* $P(e_1) = \frac{7}{18}$, *(ii)* $P(e_1) = \frac{1}{3}$, $P(e_2) = \frac{1}{6}$, and *(iii)* $P(e_1) = \frac{1}{6}$

2. *(a)* With usual notations, prove that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Deduce a similar result for $P(A \cup B \cup C)$, where $C$ is one more event.

*(b)* For any event : $E_i$, $P(E_i) = p_i$, $(i = 1, 2, 3)$ ; $P(E_i \cap E_j) = p_{ij}$, $(i, j = 1, 2, 3)$ and $P(E_1 \cap E_2 \cap E_3) = p_{123}$, find the probability that of the three events, *(i)* at least one, and *(ii)* exactly one happens.

*(c)* Discuss briefly the axiomatic approach to probability, illustrating by examples how it meets the deficiencies of the classical approach.

*(d)* If $A$ and $B$ are any two events, state the results giving

*(i)* $P(A \cup B)$ and *(ii)* $P(A \cap B)$.

$A$ and $B$ are mutually exclusive events and $P(A) = \frac{1}{2}, P(B) = \frac{1}{3}$. Find $P(A \cup B)$ and $P(A \cap B)$.

3. Let $S = \left\{1, \frac{1}{2}, \left(\frac{1}{2}\right)^2, \ldots, \left(\frac{1}{2}\right)^n\right\}$, be a classical event space and $A$, $B$ be events given by

$$A = \left\{1, \frac{1}{2}\right\}, \quad B = \left\{\left(\frac{1}{2}\right)^k \mid k \text{ is an even positive integer}\right\}$$

Find $P(\bar{A} \cap \bar{B})$ **[Calcutta Univ. B.Sc. ( Stat Hons. ), 1986]**

4. What is a 'probability space'? State *(i)* the 'law of total probability' and *(ii)* Boole's inequality for events not necessarily mutually exclusive.

5. *(a)* Explain the following with examples:

*(i)* random experiment, *(ii)* an event, *(iii)* an event space. State the axioms of probability and explain their frequency interpretations.

A man forgets the last digit of a telephone number, and dials the last digit at random. What is the probability of calling no more than three wrong numbers?

*(b)* Define conditional probability and give its frequency interpretation. Show that conditional probabilities satisfy the axioms of probability.

6. Prove the following laws, in each case assuming the conditional probabilities being defined.

*(a)* $P(E \mid E) = 1$, *(b)* $P(\phi \mid F) = 0$

*(c)* If $E_1 \subseteq E_2$, then $E(E_1 \mid F) < P(E_2 \mid F)$

*(d)* $P(\bar{E} \mid F) = 1 - P(E \mid F)$

*(e)* $P(E_1 \cup E_2 \mid F) = P(E_1 \mid F) + P(E_2 \mid F) - P(P(E_1 \cap E_2 \mid F)$

*(f)* If $P(F) = 1$ then $P(E \mid F) = P(E)$

*(g)* $P(E - F) = P(E) - P(E \cap F)$

*(h)* If $P(F) > 0$, and $E$ and $F$ are mutully exclusive then $P(E \mid F) = 0$

*(i)* If $P(E \mid F) = P(E)$, then $P(E \mid \bar{F}) = P(E)$ and $P(\bar{E} \mid F) = P(\bar{E})$

7. *(a)* If $P(\bar{A}) = a$, $P(\bar{B}) = b$, then prove that $P(A \cap B) \geq 1 - a - b$.

*(b)* If $P(A) = \alpha$, $P(B) = \beta$, then prove that $P(A \mid B) \geq (\alpha + \beta - 1)/\beta$.

**Hint.** In each case use $P(A \cup B) \leq 1$

8. Prove or disprove:

*(a) (i)* If $P(A \mid B) \geq P(A)$, then $P(B \mid A) \geq P(B)$

*(ii)* If $P(A) = P(\bar{B})$, then $A = \bar{B}$.

**[Delhi Univ. B.Sc. (Maths Hons.), 1988]**

*(b)* If $P(A) = 0$, then $A = \phi$

**[Delhi Univ. B.Sc. (Maths Hons.), 1990]**

**Ans.** Wrong.

*(c)* For possibie events $A, B, C$,

*(i)* If $P(A) > P(B)$, then $P(A \mid C) > P(B \mid C)$

*(ii)* If $P(A \mid C) \geq P(B \mid C)$ and $P(A \mid \bar{C}) \geq P(B \mid \bar{C})$, then $P(A) \geq P(B)$. **[Delhi Univ. B.Sc.(Maths Hons),1989]**

*(d)* If $P(A) = 0$, then $P(A \cap B) = 0$.

**[Delhi Univ. B.Sc. (Maths Hons.), 1986]**

*(e)* *(i)* If $P(A) = P(B) = p$, then $P(A \cap B) \leq p^2$

*(ii)* If $P(B \mid \bar{A}) = P(B \mid A)$, then $A$ and $B$ are independent.

**[Delhi Univ. B.Sc. (Maths Hons.), 1990]**

*(f)* If $P(A) > 0, P(B) > 0$ and $P(A \mid B) = P(B \mid A)$,
then $P(A) = P(B)$.

9. *(a)* Let $A$ and $B$ be two events, neither of which has probability zero. Then if $A$ and $B$ are disjoint, $A$ and $B$ are independent.

**[Delhi Univ. B.Sc.(Stat. Hons.), 1986]**

*(b)* Under what conditions does the following equality hold?

$$P(A) = P(A \mid B) + P(A \mid \overline{B})$$

**[Punjab Univ. B.Sc. (Maths Hons.), 1992]**

**Ans.** $B = S$ or $\overline{B} = S$

10. *(a)* If $A$ and $B$ are two events and the probability $P(B) \neq 1$, prove that

$$P(A \mid \overline{B}) = \frac{[P(A) - P(A \cap B)]}{[1 - P(B)]}$$

where $\overline{B}$ denotes the event complementary to $B$ and hence deduce that

$$P(A \cap B) \geq P(A) + P(B) - 1$$

**[Delhi Univ. B.Sc. (Stat. Hons.), 1989]**

Also show that $P(A) >$ *or* $< P(A \mid B)$ according as

$$P(A \mid \overline{B}) > or < P(A).$$

**[ Sri Venkat. Univ. B.Sc. 1992 ; Karnatak Univ. B.Sc.1991]**

**Hint.** *(i)*

$$P(A \mid \overline{B}) = \frac{P(A \cap \overline{B})}{P(\overline{B})} = \frac{[P(A) - P(A \cap B)]}{[1 - P(B)]}$$

*(ii)* Since $P(A \mid \overline{B}) \leq 1$, $P(A) - P(A \cap B) \leq 1 - P(B)$

$\Rightarrow \quad P(A) + P(B) - 1 \leq P(A \cap B)$

*(iii)* $$\frac{P(A \mid \overline{B})}{P(A)} = \frac{P(\overline{B} \mid A)}{P(\overline{B})} = \frac{1 - P(B \mid A)}{1 - P(B)}$$

Now $\quad P(A \mid \overline{B}) > P(A)$ if $\{1 - P(B \mid A)\} > \{1 - P(B)\}$

*i.e.*, if $\quad P(B \mid A) < P(B)$

*i.e.*, if $\quad \dfrac{P(B \mid A)}{P(B)} < 1$

*i.e.*, if $\quad \dfrac{P(A \mid B)}{P(A)} < 1$ *i.e.*, if $P(A) > P(A \mid B)$

*(b)* If $A$ and $B$ are two mutually exclusive events show that

$$P(A \mid \overline{B}) = P(A)/[1 - P(B)]$$

**[Delhi Univ. B.Sc. ( Stat. Hons.), 1987]**

*(c)* If $A$ and $B$ are two mutually exclusive events and $P(A \cup B) \neq 0$, then

$$P(A \mid A \cup B) = \frac{P(A)}{P(A) + P(B)}$$

**[Guahati Univ. B.Sc. 1991]**

*(d)* If $A$ and $B$ are two independent events show that

$$P(A \cup B) = 1 - P(\overline{A})P(\overline{B})$$

*(e)* If $\overline{A}$ denotes the non-occurrence of $A$, then prove that

$$P(\overline{A}_1 \cup \overline{A}_2 \cup \overline{A}_3) = 1 - P(A_1)\ P(A_2 \mid A_1)\ P(A_3 \mid A_1 \cap A_2)$$

**[Agra Univ. B.Sc., 1987]**

**11.** If $A, B$ and $C$ are three arbitrary events and

$$S_1 = P(A) + P(B) + P(C)$$
$$S_2 = P(A \cap B) + P(B \cap C) + P(C \cap A)$$
$$S_3 = P(A \cap B \cap C).$$

Prove that the probability that exactly one of the three events occurs is given by $S_1 - 2S_2 + 3S_3$.

**12.** *(a)* For the events $A_1, A_2, \ldots, A_n$ assuming

$$P\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} P(A_i), \text{ prove that}$$

*(i)* $P\left(\bigcap_{i=1}^{n} A_i\right) \ge 1 - \sum_{i=1}^{n} P(\overline{A}_i)$ and that

*(ii)* $P\left(\bigcap_{i=1}^{n} A_i\right) \ge \sum_{i=1}^{n} P(A_i) - (n-1)$

**[Sardar Patel Univ. B.Sc. Nov.1992]**

*(b)* Let $A, B$ and $C$ denote events. If $P(A \mid C) \ge P(B \mid C)$ and $P(A \mid \overline{C}) \ge P(B \mid \overline{C})$, then show that $P(A) \ge P(B)$.

**[Calcutta Univ. B.Sc. (Maths Hons.), 1992]**

**13.** *(a)* If $A$ and $B$ are independent events defined on a given probability space $(\Omega, \mathcal{A}, P(.))$, then prove that $A$ and $\overline{B}$ are independent, $\overline{A}$ and $\overline{B}$ are independent. **[Delhi Univ. B.Sc. (Maths Hons.), 1988]**

*(b)* $A, B$ and $C$ are three events such that $A$ and $B$ are independent, $P(C) = 0$. Show that $A, B$ and $C$ are independent.

*(c)* An event $A$ is known to be independent of the events $B, B \cup C$ and $B \cap C$. Show that it is also independent of $C$. **[Nagpur Univ. B.Sc.1992]**

*(d)* Show that if an event $C$ is independent of two mutually exclusive events $A$ and $B$, then $C$ is also independent of $A \cup B$.

*(e)* The outcome of an experiment is equally likely to be one of the four points in three-dimensional space with rectangular coordinates $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ and $(1, 1, 1)$. Let $E, F$ and $G$ be the events : $x$-coordinate=1, $y$-coordinate=1 and $z$-coordinate=1, respectively. Check if the events $E, F$ and $G$ are independent. **(Calcutta Univ. B.Sc., 1988)**

**14.** Explain what is meant by "Probability Space". You fire at a target with each of the three guns; $A, B$ and $C$ denote respectively the event — hit the target with the first, second and third gun. Assuming that the events are independent and have probabilities $P(A) = a$, $P(B) = b$ and $P(C) = c$, express in terms of $A, B$ and $C$ the following events:

*(i)* You will not hit the target at all.

(*ii*) You will hit the target at least twice. Find also the probabilities of these events. **[Sardar Patel Univ. B.Sc., 1990)**

**15.** (*a*) Suppose $A$ and $B$ are any two events and that $P(A) = p_1$, $P(B) = p_2$ and $P(A \cap B) = p_3$. Show that the formula of each of the following probabilities in terms of $p_1$, $p_2$ and $p_3$ can be expressed as follows :

(*i*) $P(\overline{A} \cup \overline{B}) = 1 - p_2$ (*ii*) $P(\overline{A} \cap \overline{B}) = 1 - p_1 - p_2 + p_3$

(*iii*) $P(A \cap \overline{B}) = p_1 - p_3$ (*iv*) $P(\overline{A} \cap B) = p_2 - p_3$

(*v*) $P(\overline{A \cap B}) = 1 - p_3$ (*vi*) $P(\overline{A} \cup B) = 1 - p_1 + p_3$

(*vii*) $P(\overline{A \cup B}) = 1 - p_1 - p_2 + p_3$ (*viii*) $P[\overline{A} \cap (A \cup B)] = p_2 - p_3$

(*ix*) $P[A \cup (\overline{A} \cap B)] = p_1 + p_2 - p_3$

(*x*) $P(A \mid B) = \dfrac{p_3}{p_2}$ and $P(B \mid A) = \dfrac{p_3}{p_1}$

(*xi*) $P(\overline{A} \mid \overline{B}) = \dfrac{1 - p_1 - p_2 + p_3}{1 - p_2}$ and $P(\overline{B} \mid \overline{A}) = \dfrac{1 - p_1 - p_2 + p_3}{1 - p_1}$

**[Allahabad Univ. B.Sc. (Stat.), 1991]**

(*b*) If $P(A) = 1/3$, $P(B) = 3/4$ and $P(A \cup B) = 11/12$, find $P(A \mid B)$ and $P(B \mid A)$.

(*c*) Let $P(A) = p$, $P(A \mid B) = q$, $P(B \mid A) = r$. Find the relation between the numbers $p$, $q$ and $r$ such that $\overline{A}$ and $\overline{B}$ are mutually exclusive.

**[Delhi Univ. B.Sc. (Maths Hons.), 1985]**

**Hint.** $P(AB) = P(A)\,P(B \mid A) = P(B) \,.\, P(A \mid B)$

$\Rightarrow \quad P(AB) = pr = P(B).\, q \quad \Rightarrow \quad P(B) = pr/q$

If $\overline{A}$ and $\overline{B}$ are mutually disjoint, then $P(\overline{A} \cap \overline{B}) = 0$.

$\Rightarrow \quad 1 - P(A \cup B) = 0 \Rightarrow 1 - [p + (pr/q) - pr] = 0$

**16.** (*a*) In terms of probabilities, $p_1 = P(A)$, $p_2 = P(B)$ and $p_3 = P(A \cap B)$;

Express (*i*) $P(A \cup B)$, (*ii*) $P(A \mid B)$, (*iii*) $P(\overline{A} \cap B)$ under the condition that (*i*) $A$ and $B$ are mutually exclusive, (*ii*) $A$ and $B$ are mutually independent.

(*b*) Let $A$ and $B$ be the possible outcomes of an experiment and suppose

$P(A) = 0{\cdot}4$, $P(A \cup B) = 0{\cdot}7$ and $P(B) = p$

(*i*) For what choice of $p$ are $A$ and $B$ mutually exclusive ?

(*ii*) For what choice of $p$ are $A$ and $B$ independent ?

**[Aligarh Univ. B.Sc., 1988 ; Guwahati Univ. B.Sc., 1991]**

**Ans.** (*i*) 0·3, (*ii*) 0·5

(*c*) Let $A_1, A_2, A_3, A_4$ be four independent events for which $P(A_1) = p$, $p(A_2) = q$, $P(A_3) = r$ and $P(A_4) = s$. Find the probability that

(*i*) at least one of the events occurs, (*ii*) exactly two of the events occur, and (*iii*) at most three of the events occur. **[Civil Services (Main), 1985]**

**17.** (*a*) Two six-faced unbiased dice are thrown. Find the probability that the sum of the numbers shown is 7 or their product is 12.

**Ans.** 2/9

*(b)* Defects are classifed as $A, B$ or $C$, and the following probabilities have been determined from available production data :

$P(A) = 0{\cdot}20, P(B) = 0{\cdot}16, P(C) = 0{\cdot}14, P(A \cap B) = 0{\cdot}08, P(A \cap C) = 0{\cdot}05, P(B \cap C) = 0{\cdot}04$, and $P(A \cap B \cap C) = 0{\cdot}02$.

What is the probability that a randomly selected item of product will exhibit at least one type of defect ? What is the probability that it exhibits both $A$ and $B$ defects but is free from type $C$ defect ? **[Bombay Univ. B.Sc., 1991]**

*(c)* A language class has only three students $A, B, C$ and they independently attend the class. The probabilities of attendance of $A, B$ and $C$ on any given day are $1/2, 2/3$ and $3/4$ respectively. Find the probability that the total number of attendances in two consecutive days is exactly three.

**[Lucknow Univ. B.Sc. 1990; Calcutta Univ. B.Sc.(Maths Hons.), 1986]**

**18.** *(a)* Cards are drawn one by one from a full deck. What is the probability that exactly 10 cards will precede the first ace. **[Delhi Univ. B.Sc.,1988]**

**Ans.** $\left(\frac{48}{52} \times \frac{47}{51} \times \frac{46}{50} \times \dots \times \frac{39}{43}\right) \times \frac{4}{42} = \frac{164}{4165}$

*(b)* Each of two persons tosses three fair coins. What is the probability that they obtain the same number of heads.

**Ans.** $\left(\frac{1}{8}\right)^2 + \left(\frac{3}{8}\right)^2 + \left(\frac{3}{8}\right)^2 + \left(\frac{1}{8}\right)^2 = \frac{5}{16}$.

**19.** *(a)* Given that $A, B$ and $C$ are mutually exclusive events, explain why each of the following is not a permissible assignment of probabilities.

*(i)* $P(A) = 0{\cdot}24$, $P(B) = 0{\cdot}4$ and $P(A \cup C) = 0{\cdot}2$,

*(ii)* $P(A) = 0{\cdot}7$, $P(B) = 0{\cdot}1$ and $P(B \cap C) = 0{\cdot}3$

*(iii)* $P(A) = 0{\cdot}6$, $P(A \cap \overline{B}) = 0{\cdot}5$

*(b)* Prove that for $n$ arbitrary independent events $A_1, A_2, \dots, A_n$

$$P(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n) + P(\overline{A_1})\, P(\overline{A_2}) \dots P(\overline{A_n}) = 1.$$

*(c)* $A_1, A_2, \dots, A_n$ are $n$ independent events with

$$P(A_i) = 1 - \frac{1}{\alpha^i},\ i = 1, 2, \dots, n.$$

Find the value of $P(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n)$. **(Nagpur Univ. B.Sc., 1987)**

**Ans.** $1 - \dfrac{1}{\alpha^{n(n+1)/2}}$

*(d)* Suppose the events $A_1, A_2, \dots, A_n$ are independent and that $P(A_i) = \dfrac{1}{i+1}$ for $1 \le i \le n$. Find the probability that none of the $n$ events occurs, justifying each step in your calculations.

**Ans.** $1/(n+1)$

**20.** *(a)* $A$ denotes getting a heart card, $B$ denotes getting a face card (King, Queem or Jack), $\overline{A}$ and $\overline{B}$ denote the complementary events. A card is drawn at random from a full deck. Compute the following probabilities.

(*i*) $P(A)$, (*ii*) $P(A \cap \overline{B})$, (*iii*) $P(A \cup B)$, (*iv*) $P(A \cap B)$,

(*v*) $P(\overline{A} \cup B)$.

Assume natural assignment of probabilities.

**Ans.** (*i*) 1/4, (*ii*) 5/26, (*iii*) 11/26, (*iv*) 3/5, (*v*) 21/26.

(*b*) A town has two doctors $X$ and $Y$ operating independently. If the probability that doctor $X$ is available is 0·9 and that for $Y$ is 0·8, what is the probability that at least one doctor is available when needed? **[Gorakhpur Univ. B.Sc., 1988]**

**Ans.** 0·98

**21.** (*a*) The odds that a book will be favourably reviewed by 3 independent critics are 5 to 2, 4 to 3 and 3 to 4 respectively. What is the probability that, of the three reviews, a majority will be favourable? **[Gauhati Univ. BSc., 1987]**

**Ans.** 209/343.

(*b*) $A$, $B$ and $C$ are independent witnesses of an event which is known to have occurred. $A$ speaks the truth three times out of four, $B$ four times out of five and $C$ five times out of six. What is the probability that the occurrence will be reported truthfully by majority of three witnesses?

**Ans.** 31/60.

(*c*) A man seeks advice regarding one of two possible courses of action from three advisers who arrived at their recommendations independently. He follows the recommendation of the majority. The probability that the individual advisers are wrong are 0·1, 0·05 and 0·05 respectively. What is the probability that the man takes incorrect advise ? **[Gujarat Univ. B.Sc., 1987]**

**22.** (*a*) The odds against a certain event are 5 to 2 and odds in favour of another (independent) event are 6 to 5. Find the chance that at least one of the events will happen. **(Madras Univ. B.Sc.,1987)**

**Ans.** 52/77.

(*b*) A person takes four tests in succession. The probability of his passing the first test is $p$, that of his passing each succeeding test is $p$ or $p/2$ according as he passes or fails the preceding one. He qualifies provided he passes at least three tests. What is his chance of qualifying. **[Gauhati Univ. B.Sc. (Hons.) 1988]**

**23.** (*a*) The probability that a 50-years old man will be alive at 60 is 0·83 and the probability that a 45-years old woman will be alive at 55 is 0·87. What is the probability that a man who is 50 and his wife who is 45 will both be alive 10 years hence?

**Ans.** 0·7221.

(*b*) It is 8:5 against a husband who is 55 years old living till he is 75 and 4:3 against his wife who is now 48, living till she is 68. Find the probability that (*i*) the couple will be alive 20 years hence, and (*ii*) at least one of them will be alive 20 years hence.

**Ans** (*i*) 15/91, (*ii*) 59/91.

*(c)* A husband and wife appear in an interview for two vacancies in the same post. The probability of husband's selection is 1/7 and that of wife's selection is 1/5. What is the probability that only one of them will be selected ?

**Ans.** 2/7 **[Delhi Univ. B.Sc., 1986]**

**24.** *(a)* The chances of winning of two race-horses are 1/3 and 1/6 respectively. What is the probability that at least one will win when the horses are running *(a)* in different races, and *(b)* in the same race?

**Ans.** *(a)* 8/18 *(b)* 1/2

*(b)* A problem in statistics is given to three students whose chances of solving it are 1/2, 1/3 and 1/4. What is the probability that the problem will be solved?

**Ans.** 3/4 **[Meerut Univ. B.Sc., 1990]**

**25.** *(a)* Ten pairs of shoes are in a closet. Four shoes are selected at random. Find the probability that there will be at least one pair among the four shoes selected?

**Ans.** $1 - \dfrac{{}^{10}C_4 \times 2^4}{{}^{20}C_4}$

*(b)* From 100 tickets numbered 1, 2, ..., 100 four are drawn at random. What is the probability that 3 of them will bear number from 1 to 20 and the fourth will bear any number from 21 to 100 ?

**Ans.** $\dfrac{{}^{20}C_3 \times {}^{80}C_1}{{}^{100}C_4}$

**26.** A six faced die is so biased that it is twice as likely to show an even number as an odd number when thrown. It is thrown twice. What is the probability that the sum of the two numbers thrown is odd?

**Ans.** 4/9

**27.** From a group of 8 children, 5 boys and 3 girls, three children are selected at random. Calculate the probabilities that selected group contains *(i)* no girl, *(ii)* only one girl, *(iii)* one particular girl, *(iv)* at least one girl, and *(v)* more girls than boys.

**Ans.** *(i)* 5/28, *(ii)* 15/28, *(iii)* 5/28, *(iv)* 23/28, *(v)* 2/7.

**28.** If three persons, selected at random, are stopped on a street, what are the probabilities that :

*(a)* all were born on a Friday;

*(b)* two were born on a Friday and the other on a Tuesday;

*(c)* none was born on a Monday.

**Ans.** *(a)* 1/343, *(b)* 3/343, *(c)* 216/343.

**29.** *(a)* *A* and *B* toss a coin alternately on the understanding that the first who obtains the head wins. If *A* starts, show that their respective chances of winning are 2/3 and 1/3.

*(b)* *A*, *B* and *C*, in order, toss a coin. The first one who throws a head wins. If *A* starts, find their respectivre chances of winning. (Assume that the game may

continue indefinitely.)

**Ans.** 4/7, 2/7, 1/7.

*(c)* A man alternately tosses a coin and throws a die, beginning with coin. What is the probability that he will get a head before he gets a '5 or 6' on die?

**Ans.** 3/4.

**30.** *(a)* Two ordinary six-sided dice are tossed.

*(i)* What is the probability that both the dice show the number 5.

*(ii)* What is the probability that both the dice show the same number.

*(iii)* Given that the sum of two numbers shown is 8, find the conditional probability that the number noted on the first dice is larger than the number noted on the second dice.

*(b)* Six dice are thrown simultaneously. What is the probability that all will show different faces?

**31.** *(a)* A bag contains 10 balls, two of which are red, three blue and five black. Three balls are drawn at random from the bag, that is every ball has an equal chance of being included in the three. What is the probability that

*(i)* the three balls are of different colours,

*(ii)* two balls are of the same colour, and

*(iii)* the balls are all of the same colour?

**Ans.** *(i)* 30/120, *(ii)* 79/120, *(iii)* 11/120.

*(b)* $A$ is one of six horses entered for a race and is to be ridden by one of the two jockeys $B$ and $C$. It is 2 to 1 that $B$ rides $A$, in which case all the horses are equally likely to win, with rider $C$, $A$'s chance is trebled.

*(i)* Find the probability that $A$ wins.

*(ii)* What are odds against $A$'s winning?

**[Shivaji Univ. B.Sc. (Stat. Hons.), 1992]**

**Hint.** Probability of $A$'s winning

$$= P(B \text{ rides } A \text{ and } A \text{ wins}) + P(C \text{ rides } A \text{ and } A \text{ wins})$$

$$= \frac{2}{3} \times \frac{1}{6} + \frac{1}{3} \times \frac{3}{6} = \frac{5}{18}$$

$\therefore$ Probability of $A$'s losing $= 1 - 5/18 = 13/18$.

Hence odds against $A$'s winning are : $13/18 : 5/18$, *i.e.*, $13 : 5$.

**32.** *(a)* Two-third of the students in a class are boys and the rest girls. It is known that the probability of a girl getting a first class is 0·25 and that of boy getting a first class is 0·28. Find the probability that a student chosen at random will get first class marks in the subject.

**Ans.** 0·27

*(b)* You need four eggs to make omelettes for breakfast. You find a dozen eggs in the refrigerator but do not realise that two of these are rotten. What is the probability that of the four eggs you choose at random

*(i)* none is rotten,

*(ii)* exactly one is rotten?

**Ans.** *(i)* 625/1296 ; *(ii)* 500/1296.

*(c)* The probability of occurrence of an event $A$ is 0·7, the probability of non-occurrence of another event $B$ is 0·5 and that of at least one of $A$ or $B$ not occurring is 0·6. Find the probability that at least one of $A$ or $B$ occurs.

**[Mysore Univ. B.Sc., 1991]**

**33.** *(a)* The odds against $A$ solving a certain problem are 4 to 3 and odds in favour of $B$ solving the same problem are 7 to 5. What is the probability that the problem is solved if they both try independently? **[Gujarat Univ.B.Sc., 1987]**

**Ans.** 16/21

*(b)* A certain drug manufactured by a company is tested chemically for its toxic nature. Let the event 'the drug is toxic' be denoted by $E$ and the event 'the chemical test reveals that the drug is toxic' be denoted by $F$. Let $P(E) = \theta$, $P(F \mid E) = P(\bar{F} \mid \bar{E}) = 1 - \theta$. Then show that probability that the drug is not toxic given that the chemical test reveals that it is toxic is free from $\theta$.

**Ans.** 1/2 **[M.S. Baroda Univ. B.Sc., 1992]**

**34.** A bag contains 6 white and 9 black balls. Four balls are drawn at a time. Find the probability for the first draw to give 4 white and the second draw to give 4 black balls in each of the following cases :

*(i)* The balls are replaced before the second draw.

*(ii)* The balls are not replaced before the second draw.

**[Jammu Univ. B.Sc., 1992]**

**Ans.** *(i)* $\dfrac{{}^6C_4}{{}^{15}C_4} \times \dfrac{{}^9C_4}{{}^{15}C_4}$ *(ii)* $\dfrac{{}^6C_4}{{}^{15}C_4} \times \dfrac{{}^9C_4}{{}^{11}C_4}$

**35.** The chances that doctor $A$ will diagnose a disease $X$ correctly is 60%. The chances that a patient will die by his treatment after correct diagnosis is 40% and the chance of death by wrong diagnosis is 70%. A patient of doctor $A$, who had disease $X$, died. What is the chance that his disease was diagnosed correctly?

**Hint.** Let us define the following events:

$E_1$ : Disease $X$ is diagnosed correctly by doctor $A$.

$E_2$ : A patient (of doctor $A$) who has disease $X$ dies.

Then we are given :

$$P(E_1) = 0{\cdot}6 \quad \Rightarrow \quad P(\bar{E}_1) = 1 - 0{\cdot}6 = 0{\cdot}4$$

$$\text{and} \quad P(E_2 \mid E_1) = 0{\cdot}4 \quad \text{and} \quad P(E_2 \mid \bar{E}_1) = 0{\cdot}7$$

We want $P(E_1 \mid E_2) = \dfrac{P(E_1 \cap E_2)}{P(E_2)} = \dfrac{P(E_1 \cap E_2)}{P(E_1 \cap E_2) + P(\bar{E}_1 \cap E_2)} = \dfrac{6}{13}$

**36.** The probability that at least 2 of 3 people $A$, $B$ and $C$ will survive for 10 years is 247/315. The probability that $A$ alone will survive for 10 years is 4/105 and the probability that $C$ alone will die within 10 years is 2/21. Assuming that the events of the survival of $A$, $B$ and $C$ can be regarded as independent, calculate the probability of surviving 10 years for each person.

Ans. 3/5, 5/7, 7/9.

37. $A$ and $B$ throw alternately a pair of unbiased dice, $A$ beginning. $A$ wins if he throws 7 before $B$ throws 6, and $B$ wins if he throws 6 before $A$ throws 7. If $A$ and $B$ respectively denote the events that $A$ wins and $B$ wins the series, and $a$ and $b$ respectively denote the events that it is $A$'s and $B$'s turn to throw the dice, show that

$$(i)\ P(A \mid a) = \frac{1}{6} + \frac{5}{6} P(A \mid b), \quad (ii)\ P(A \mid b) = \frac{31}{36} P(A \mid a),$$

$$(iii)\ P(B \mid a) = \frac{5}{6} P(B \mid b), \text{ and } (iv)\ P(B \mid b) = \frac{5}{36} + \frac{13}{36} P(B \mid a),$$

Hence or otherwise, find $P(A \mid a)$ and $P(B \mid a)$. Also comment on the result that $P(A \mid a) + P(B \mid a) = 1$.

38. A bag contains an assortment of blue and red balls. If two balls are drawn at randon, the probability of drawing two red balls is five times the probability of drawing two blue balls. Furthermore, the probability of drawing one ball of each colour is six times the probabililty of drawing two blue balls. How many red and blue balls are there in the bag?

**Hint.** Let number of red and blue balls in the bag be $r$ and $b$ respectively. Then

$$p_1 = \text{ Prob. of drawing two red balls } = \frac{r(r-1)}{(r+b)(r+b-1)}$$

$$p_2 = \text{ Prob. of drawing two blue balls} = \frac{b(b-1)}{(r+b)(r+b-1)}$$

$$p_3 = \text{ Prob. of drawing one red and one blue ball } = \left[\frac{2br}{(r+b)(r+b-1)}\right].$$

Now $p_1 = 5p_2$ and $p_3 = 6p_2$

$\therefore$ $r(r-1) = 5b(b-1)$ and $2br = 6b(b-1)$

Hence $b = 3$ and $r = 6$.

39. Three newspapers $A$, $B$ and $C$ are published in a certain city. It is estimated from a survey that 20% read $A$; 16% read $B$, 14% read $C$, 8% read $A$ and $B$, 5% read $A$ and $C$, 4% read $B$ and $C$ and 2% read all the three newspapers. What is the probability that a normally chosen person

*(i)* does not read any paper, *(ii)* does not read $C$

*(iii)* reads $A$ but not $B$, *(iv)* reads only one of these papers, and

*(v)* reads only two of these papers.

Ans. *(i)* 0·65, *(ii)* 0·86, *(iii)* 0·12, *(iv)* 0·22, *(v)* 0·11.

40. *(a)* A die is thrown twice, the event space $S$ consisting of the 36 possible pairs of outcomes $(a,b)$ each assigned probability 1/36. Let $A$, $B$ and $C$ denote the following events :

$A = \{(a,b) \mid a \text{ is odd}\}$, $B = \{(a,b) \mid b \text{ is odd}\}$, $C = \{(a,b) \mid a + b \text{ is odd.}\}$

Check whether $A$, $B$ and $C$ are independent or independent in pairs only.

[Calcutta Univ. B.Sc. Hons., 1985]

*(b)* Eight tickets numbered 111, 121, 122, 122, 211, 212, 212, 221 are placed in a hat and stirred. One of them is then drawn at random. Show that the event $A$ : "the first digit on the ticket drawn will be 1", $B$ : "the second digit on the ticket drawn will be 1," and $C$ : "the third digit on the ticket drawn will be 1", are not pairwise independent although

$$P(A \cap B \cap C) = P(A)\,P(B)\,P(C)$$

**41.** *(a)* Four identical marbles marked 1, 2, 3 and 123 respectively are put in an urn and one is drawn at random. Let $A_i$, $(i = 1, 2, 3)$, denote the event that the number $i$ appears on the drawn marble. Prove that the events $A_1, A_2$ and $A_3$ are pairwise independent but not mutually independent.

**[Gauhati Univ. B.Sc. (Hons.), 1988]**

**Hint.** $P(A_1) = \frac{1}{2} = P(A_2) = P(A_3)$; $P(A_1 A_2) = P(A_1 A_3) = P(A_2 A_3) = \frac{1}{4}$

$P(A_1 A_2 A_3) = \frac{1}{4}$.

*(b)* Two fair dice are thrown independently. Define the following events :

$A$ : Even number on the first dice

$B$ : Even number on the second dice.

$C$ : Same number on both dice.

Discuss the independence of the events $A$, $B$ and $C$.

*(c)* A die is of the shape of a regular tetrahedron whose faces bear the numbers 111, 112, 121, 122. $A_1, A_2, A_3$ are respectively the events that the first two, the last two and the extreme two digits are the same, when the die is tossed at random. Find whether or not the events $A_1, A_2, A_3$ are *(i)* pairwise independent, *(ii)* mutually (*i.e.* completely) independent. Determine $P(A_1 \mid A_2 A_3))$ and explain its value by by argument. **[Civil Services (Main), 1983]**

**42.** *(a)* For two events $A$ and $B$ we have the following probabilities:

$$P(A) = P(A \mid B) = \frac{1}{4} \text{ and } P(B \mid A) = \frac{1}{2}.$$

Check whether the following statements are true or false :

*(i)* $A$ and $B$ are mutually exclusive, *(ii)* $A$ and $B$ are independent, *(iii)* $A$ is a subevent of $B$, and *(iv)* $P(\bar{A} \mid B) = \frac{3}{4}$

**Ans.** *(i)* False, *(ii)* True, *(iii)* False, and *(iv)* True.

*(b)* Consider two events $A$ and $B$ such that $P(A) = 1/4$, $P(B \mid A) = 1/2$, $P(A \mid B) = 1/4$. For each of the following statements, ascertain whether it is true or false :

*(i)* $A$ is a sub-event of $B$, *(ii)* $P(\bar{A} \mid \bar{B}) = 3/4$,

*(iii)* $P(A \mid B) + P(A \mid \bar{B}) = 1$

**43.** *(a)* Let $A$ and $B$ be two events such that $P(A) = \frac{3}{4}$ and $P(B) = \frac{5}{8}$.

Show that

(i) $P(A \cup B) \geq \frac{3}{4}$, (ii) $\frac{3}{8} \leq P(A \cap B) \leq \frac{5}{8}$, and (iii) $\frac{1}{8} \leq P(A \cap \bar{B}) \leq \frac{3}{8}$.

**[Coimbatore Univ. B.E., Nov. 1990; Delhi Univ. B.Sc.(Stat. Hons.),1986]**

*(b)* Given two events $A$ and $B$. If the odds against $A$ are 2 to 1 and those in favour of $A \cup B$ are 3 to 1, show that

$$\frac{5}{12} \leq P(B) \leq \frac{3}{4}$$

Give an example in which $P(B) = 3/4$ and one in which $P(B) = 5/12$.

**44.** Let $A$ and $B$ be events, neither of which has probability zero. Prove or disprove the following events :

*(i)* If $A$ and $B$ are disjoint, $A$ and $B$ are independent,

*(ii)* If $A$ and $B$ are independent, $A$ and $B$ are disjoint.

**45.** *(a)* It is given that $P(A_1 \cup A_2) = \frac{5}{6}$, $P(A_1 \cap A_2) = \frac{1}{3}$ and $P(\bar{A}_2) = \frac{1}{2}$, where $P(\bar{A}_2)$ stands for the probability that $A_2$ does not happen. Determine $P(A_1)$ and $P(A_2)$.

Hence show that $A_1$ and $A_2$ are independent.

**Ans.** $P(A_1) = \frac{2}{3}$, $P(A_2) = \frac{1}{2}$

*(b)* $A$ and $B$ are events such that

$$P(A \cup B) = \frac{3}{4},\ P(A \cap B) = \frac{1}{4},\ \text{and}\ P(\bar{A}) = \frac{2}{3}.$$

Find *(i)* $P(A)$, *(ii)* $P(B)$ and *(iii)* $P(A \cap \bar{B})$.

**(Madras Univ. B.E., 1989)**

**Ans.** *(i)* 1/3, *(ii)* 2/3 *(iii)* 1/12.

**46.** A thief has a bunch of $n$ keys, exactly one of which fits a lock. If the thief tries to open the lock by trying the keys at random, what is the probability that he requires exactly $k$ attempts, if he rejects the keys already tried? Find the same probability if he does not reject the keys already tried.

**(Aligarh Univ. B.Sc., 1991)**

**Ans.** *(i)* $\frac{1}{n}$, *(ii)* $\left(\frac{n-1}{n}\right)^{k-1} \cdot \frac{1}{n}$

*(b)* There are $M$ urns numbered 1 to $M$ and $M$ balls numbered 1 to $M$. The balls are inserted randomly in the urns with one ball in each urn. If a ball is put into the urn bearing the same number as the ball, a match is said to have occurred. Find the probability that no match has occurred. **[Civil Services (Main), 1984]**

**Hint.** See Example 4·54 page 4·97.

**47.** If $n$ letters are placed at random in $n$ correctly addressed envelopes, find the probability that

*(i)* none of the letters is placed in the correct envelope,

*(ii)* At least one letter goes to the correct envelope,

*(iii)* All letters go to the correct envelopes.

**[Delhi Univ. B.Sc. (Stat Hons.), 1987, 1984]**

48. An urn contains $n$ white and $m$ black balls, a second urn contains $N$ white and $M$ black balls. A ball is randomly transferred from the first to the second urn and then from the second to the first urn. If a ball is now selected randomly from the first urn, prove that the probability that it is white is

$$\frac{n}{n+m}+\frac{mN-nM}{(n+m)^2\,(N+M+1)}$$

**[Delhi Univ. B.Sc. (Stat.Hons.) 1986]**

**Hint.** Let us define the following events :

$B_i$ : Drawing of a black ball from the $i$th urn, $i = 1, 2$.

$W_i$ : Drawing of a white ball from the $i$th urn , $i = 1, 2$.

The four distinct possibilities for the first two exchanges are $B_1 W_2, B_1 B_2, W_1 B_2, W_1 W_2$. Hence if $E$ denotes the event of drawing a white ball from the first urn after the exchanges, then

$$P(E) = P(B_1 W_2 E) + P(B_1 B_2 E) + P(W_1 B_2 E) + P(W_1 W_2 E) \qquad \ldots(*)$$

We have :

$$P(B_1 W_2 E) = P(B_1)\,.\,P(W_2 \mid B_1)\,P(E \mid B_1 W_2) = \frac{m}{m+n}\times\frac{N}{M+N+1}\times\frac{n+1}{m+n}$$

$$P(B_1 B_2 E) = P(B_1)\,.\,P(B_2 \mid B_1)\,.\,P(E \mid B_1 B_2) = \frac{m}{m+n}\times\frac{M+1}{M+N+1}\times\frac{n}{m+n}$$

$$P(W_1 B_2 E) = P(W_1)\,.\,P(B_2 \mid W_1)\,.\,P(E \mid W_1 B_2) = \frac{n}{m+n}\times\frac{M}{M+N+1}\times\frac{n-1}{m+n}$$

$$P(W_1 W_2 E) = P(W_1)\,.\,P(W_2 \mid W_1)\,.\,P(E \mid W_1 W_2) = \frac{n}{m+n}\times\frac{N+1}{M+N+1}\times\frac{n}{m+n}$$

Substituting in (*) and simplifying we get the result.

49. A particular machine is prone to three similar types of faults $A_1, A_2$ and $A_3$. Past records on breakdowns of the machine show the following : the probability of a breakdown (*i.e.*, at least one fault) is 0·1; for each $i$, the probability that fault $A_i$ occurs and the others do not is 0·02 ; for each pair $i, j$ the probability that $A_i$ and $A_j$ occur but the third fault does not is 0·012. Determine the probabilities of

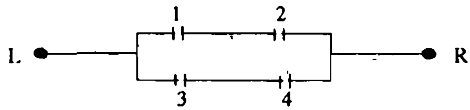
*(a)* the fault of type $A_1$ occurring irrespective of whether the other faults occur or not,

*(b)* a fault of type $A_1$ given that $A_2$ has occurred,

*(c)* faults of type $A_1$ and $A_2$ given that $A_3$ has occurred.

**[London U. B.Sc. 1976]**

50. The probability of the closing of each relay of the circuit shown below is given by $p$. If all the relays function independently, what is the probability that a circuit exists between the terminals L and R?




Ans. $p^2(2-p^2)$.

**4·9. Bayes Theorem.** *If $E_1, E_2, ..., E_n$ are mutually disjoint events with $P(E_i) \neq 0$, $(i = 1, 2, \ldots, n)$ then for any orbitrary event A which is a subset of $\bigcup_{i=1}^{n} E_i$ such that $P(A) > 0$, we have*

$$P(E_i \mid A) = \frac{P(E_i)\, P(A \mid E_i)}{\sum_{i=1}^{n} P(E_i)\, P(A \mid E_i)}, \quad i = 1, 2, ..., n. \qquad ...(4\cdot12)$$

**Proof.** Since $A \subset \bigcup_{i=1}^{n} E_i$, we have

$$A = A \cap \left(\bigcup_{i=1}^{n} E_i\right) = \bigcup_{i=1}^{n} (A \cap E_i) \qquad \text{[By distributive law]}$$

Since $(A \cap E_i) \subset E_i$, $(i = 1, 2, ...., n)$ are mutually disjoint events, we have by addition theorem of probability (or Axiom 3 of probability)

$$P(A) = P\left[\bigcup_{i=1}^{n} (A \cap E_i)\right] = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(E_i)\, P(A \mid E_i), \qquad ...(*)$$

by compound theorem of probability.

Also we have

$$P(A \cap E_i) = P(A)\, P(E_i \mid A)$$

$$P(E_i \mid A) = \frac{P(A \cap E_i)}{P(A)} = \frac{P(E_i)\, P(A \mid E_i)}{\sum_{i=1}^{n} P(E_i)\, P(A \mid E_i)} \qquad \text{[From (*)]}$$

**Remarks. 1.** The probabilities $P(E_1), P(E_2), ..., P(E_n)$ are termed as the '*a priori probabilities*' because they exist before we gain any information from the experiment itself.

2. The probabilities $P(A \mid E_i)$, $i = 1, 2, ..., n$ are called '*likelihoods*' because they indicate how likely the event $A$ under consideration is to occur, given each and every *a priori* probability.

3. The probabilities $P(E_i \mid A)$, $i = 1, 2, ..., n$ are called '*posterior probabilities*' because they are determined after the results of the experiment are known.

4. From (*) we get the following important result:

"If the events $E_1, E_2, ..., E_n$ constitute a partition of the sample space $S$ and $P(E_i) \neq 0$, $i = 1, 2, \ldots, n$, then for any event $A$ in $S$ we have

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(E_i)\, P(A \mid E_i) \quad \text{...(4·12 } a)$$

**Cor.** *(Bayes theorem for future events)*

*The probability of the materialisation of another event C, given*

$$P(C \mid A \cap E_1),\, P(C \mid A \cap E_2),\, \ldots,\, P(C \mid A \cap E_n) \quad \textit{is}$$

$$P(C \mid A) = \frac{\sum_{i=1}^{n} P(E_i)\, P(A \mid E_i)\, P(C \mid E_i \cap A)}{\sum_{i=1}^{n} P(E_i)\, P(A \mid E_i)} \quad \text{...(4·12 } b)$$

**Proof.** Since the occurrence of event $A$ implies the occurrence of one and only one of the events $E_1, E_2, \ldots, E_n$, the event $C$ (granted that $A$ has occurred) can occu in the following mutually exclusive ways:

$$C \cap E_1, C \cap E_2, \ldots, C \cap E_n$$

*i.e.,* $\quad C = (C \cap E_1) \cup (C \cap E_2) \cup \ldots \cup (C \cap E_n)$

$\Rightarrow \quad C \mid A = [(C \cap E_1) \mid A] \cup [(C \cap E_2) \mid A] \cup \ldots \cup [(C \cap E_n) \mid A]$

$\therefore \quad P(C \mid A) = P[(C \cap E_1) \mid A] + P[(C \cap E_2) \mid A] + \ldots + P[(C \cap E_n) \mid A]$

$$= \sum_{i=1}^{n} P[(C \cap E_i) \mid A]$$

$$= \sum_{i=1}^{n} P(E_i \mid A)\; P[C \mid (E_i \cap A)]$$

Substituting the value of $P(E_i \mid A)$ from (*), we get

$$P(C \mid A) = \frac{\sum_{i=1}^{n} P(E_i)\, P(A \mid E_i)\; P(C \mid E_i \cap A)}{\sum_{i=1}^{n} P(E_i)\, P(A \mid E_i)}$$

**Remark.** It may happen that the materialisation of the event $E_i$ makes $C$ independent of $A$, then we have

$$P(C \mid E_i \cap A) = P(C \mid E_i),$$

and the above formula reduces to

$$P(C \mid A) = \frac{\sum_{i=1}^{n} P(E_i)\, P(A \mid E_i)\; P(C \mid E_i)}{\sum_{i=1}^{n} P(E_i)\, P(A \mid E_i)} \quad \text{...(4·12 } c)$$

The event $C$ can be considered in regard to $A$, as Future Event.

**Example 4·30.** *In 1989 there were three candidates for the position of principal – Mr. Chatterji, Mr. Ayangar and Dr. Singh – whose chances of getting the appointment are in the proportion 4:2:3 respectively. The probability that Mr. Chatterji if selected would introduce co-education in the college is 0·3. The probabilities of Mr. Ayangar and Dr. Singh doing the same are respectively 0·5 and 0·8. What is the probability that there was co-education in the college in 1990?*

**[Delhi Univ. B.Sc.(Stat. Hons.), 1992; Gorakhpur Univ. B.Sc., 1992]**

**Solution.** Let the events and probabilities be defined as follows:

$A$ : Introduction of co-education

$E_1$ : Mr. Chatterji is selected as principal

$E_2$ : Mr. Ayangar is selected as principal

$E_3$ : Dr. Singh is selected as principal.

Then

$$P(E_1)=\frac{4}{9},\quad P(E_2)=\frac{2}{9}\quad\text{and}\quad P(E_3)=\frac{3}{9}$$

$$P(A\mid E_1)=\frac{3}{10},\quad P(A\mid E_2)=\frac{5}{10}\quad\text{and}\quad P(A\mid E_3)=\frac{8}{10}$$

$$\begin{aligned}\therefore\quad P(A) &= P[(A\cap E_1)\cup(A\cap E_2)\cup(A\cap E_3)]\\ &= P(A\cap E_1)+P(A\cap E_2)+P(A\cap E_3)\\ &= P(E_1)P(A\mid E_1)+P(E_2)P(A\mid E_2)+P(E_3)P(A\mid E_3)\\ &= \frac{4}{9}\cdot\frac{3}{10}+\frac{2}{9}\cdot\frac{5}{10}+\frac{3}{9}\cdot\frac{8}{10}=\frac{23}{45}\end{aligned}$$

**Example. 4·31.** *The contents of urns I, II and III are as follows:*

*1 white, 2 black and 3 red balls,*

*2 white, 1 black and 1 red balls, and*

*4 white, 5 black and 3 red balls.*

*One urn is chosen at random and two balls drawn. They happen to be white and red. What is the probability that they come from urns I, II or III ?*

**[Delhi Univ. B.Sc. (Stat. Hons.), 1988]**

**Solution.** Let $E_1$, $E_2$, and $E_3$ denote the events that the urn I, II and III is chosen, respectively, and let $A$ be the event that the two balls taken from the selected urn are white and red. Then

$$P(E_1)=P(E_2)=P(E_3)=\frac{1}{3}$$

$$P(A\mid E_1)=\frac{1\times 3}{{}^6C_2}=\frac{1}{5},\quad P(A\mid E_2)=\frac{2\times 1}{{}^4C_2}=\frac{1}{3},$$

and
$$P(A\mid E_3)=\frac{4\times 3}{{}^{12}C_2}=\frac{2}{11}$$

Hence

$$P(E_2 \mid A) = \frac{P(E_2)\, P(A \mid E_2)}{\sum_{i=1}^{3} P(E_i)\, P(A \mid E_i)}$$

$$= \frac{\frac{1}{3} \times \frac{1}{3}}{\frac{1}{3} \times \frac{1}{5} + \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{2}{11}} = \frac{55}{118}$$

Similarly

$$P(E_3 \mid A) = \frac{\frac{1}{3} \times \frac{2}{11}}{\frac{1}{3} \times \frac{1}{5} + \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{1}{11}} = \frac{30}{118}$$

$$\therefore \quad P(E_1 \mid A) = 1 - \frac{55}{118} - \frac{30}{118} = \frac{33}{118}$$

**Example. 4·32.** *In answering a question on a multiple choice test a student either knows the answer or he guesses. Let p be the probability that he knows the answer and 1–p the probability that he guesses. Assume that a student who guesses at the answer will be correct with probability 1/5, where 5 is the number of multiple-choice alternatives. What is the conditional probability that a student knew the answer to a question given that he answered it correctly?*

**[Delhi Univ. B.Sc. (Maths Hons.), 1985]**

**Solution.** Let us define the following events:

$E_1$ : The student knew the right answer.

$E_2$ : The student guesses the right answer.

$A$ : The student gets the right answer.

Then we are given

$P(E_1) = p, \quad P(E_2) = 1 - p, \quad P(A \mid E_2) = 1/5$

$P(A \mid E_1) = P$ [student gets the right answer given that he knew the right answer] $= 1$

We want $P(E_1 \mid A)$.

Using Bayes' rule, we get :

$$P(E_1 \mid A) = \frac{P(E_1)\, P(A \mid E_1)}{P(E_1) P(A \mid E_1) + P(E_2) P(A \mid E_2)} = \frac{p \times 1}{p \times 1 + (1-p) \times \frac{1}{5}} = \frac{5p}{4p+1}$$

**Example 4·33.** *In a bolt factory machines A, B and C manufacture respectively 25%, 35% and 40% of the total. Of their output 5, 4, 2 per cent are defective bolts. A bolt is drawn at random from the product and is found to be defective. What are the probabilities that it was manufactured by machines A, B and C?*

**Solution.** Let $E_1, E_2$ and $E_3$ denote the events that a bolt selected at random is manufactured by the machines $A, B$ and $C$ respectively and let $E$ denote the event of its being defective. Then we have

$$P(E_1) = 0{\cdot}25,\ P(E_2) = 0{\cdot}35,\ P(E_3) = 0{\cdot}40$$

The probability of drawing a defective bolt manufactured by machine $A$ is $P(E \mid E_1) = 0{\cdot}05$.

Similarly, we have

$$P(E \mid E_2) = 0{\cdot}04, \text{ and } P(E \mid E_3) = 0{\cdot}02$$

Hence the probability that a defective bolt selected at random is manufactured by machine $A$ is given by

$$P(E_1 \mid E) = \frac{P(E_1)\,P(E \mid E_1)}{\sum\limits_{i=1}^{3} P(E_i)\,P(E \mid E_i)}$$

$$= \frac{0{\cdot}25 \times 0{\cdot}05}{0{\cdot}25 \times 0{\cdot}05 + 0{\cdot}35 \times 0{\cdot}04 + 0{\cdot}40 \times 0{\cdot}02} = \frac{125}{345} = \frac{25}{69}$$

Similarly

$$P(E_2 \mid E) = \frac{0{\cdot}35 \times 0{\cdot}04}{0{\cdot}25 \times 0{\cdot}05 + 0{\cdot}35 \times 0{\cdot}04 + 0{\cdot}40 \times 0{\cdot}02} = \frac{140}{345} = \frac{28}{69}$$

and

$$P(E_3 \mid E) = 1 - [P(E_1 \mid E) + P(E_2 \mid E)] = 1 - \frac{25}{69} - \frac{28}{69} = \frac{16}{69}$$

This example illustrates one of the chief applications of Bayes Theorem.

## EXERCISE 4 (d)

1. *(a)* State and prove Baye's Theorem.

*(b)* The set of events $A_k$, $(k = 1, 2, \ldots, n)$ are *(i)* exhaustive and *(ii)* pairwise mutually exclusive. If for all $k$ the probabilities $P(A_k)$ and $P(E \mid A_k)$ are known, calculate $P(A_k \mid E)$, where $E$ is an arbitrary event. Indicate where conditions *(i)* and *(ii)* are used.

*(c)* The events $E_1, E_2, \ldots, E_n$ are mutually exclusive and $E = E_1 \cup E_2 \cup \ldots \cup E_n$. Show that if $P(A \mid E_i) = P(B \mid E_i)$ ; $i = 1, 2, \ldots, n$, then $P(A \mid E) = P(B \mid E)$. Is this conclusion true if the events $E_i$ are not mutually exclusive?

**[Calcutta Univ. B.Sc. (Maths Hons.), 1990]**

*(d)* What are the criticisms against the use of Bayes theorem in probability theory. **[Sri. Venketeswara Univ. B.Sc., 1991]**

*(e)* Using the fundamental addition and multiplication rules of probability, show that

$$P(B \mid A) = \frac{P(B)\,P(A \mid B)}{P(B)\,P(A \mid B) + P(\bar{B})\,P(A \mid \bar{B})}$$

where $\bar{B}$ is the event complementary to the event $B$.

**[Delhi Univ. M.A. (Econ.), 1987]**

2. *(a)* Two groups are competing for the positions on the Board of Directors of a corporation. The probabilities that the first and second groups will win are 0·6 and 0·4 respectively. Furthermore, if the first group wins the probability of introducing a new product is 0·8 and the corresponding probability if the second group wins is 0·3. What is the probability that the new product will be introduced?

**Ans.** $0{\cdot}6 \times 0{\cdot}8 + 0{\cdot}4 \times 0{\cdot}3 = 0{\cdot}6$

*(b)* The chances of $X, Y, Z$ becoming managers of a certain company are 4:2:3. The probabilities that bonus scheme will be introduced if $X, Y, Z$ become managers, are 0·3, 0·5 and 0·8 respectively. If the bonus scheme has been introduced, what is the probability that $X$ is appointed as the manager.

**Ans.** 0·51

*(c)* A restaurant serves two special dishes, $A$ and $B$ to its customers consisting of 60% men and 40% women. 80% of men order dish $A$ and the rest $B$. 70% of women order dish $B$ and the rest $A$. In what ratio of $A$ to $B$ should the restaurant prepare the two dishes? **(Bangalore Univ. B.Sc., 1991)**

**Ans.** $P(A) = P[(A \cap M) \cup (A \cap W)] = 0{\cdot}6 \times 0{\cdot}8 + 0{\cdot}4 \times 0{\cdot}3 = 0{\cdot}6$

Similarly $P(B) = 0{\cdot}4$. Required ratio $= 0{\cdot}6 : 0{\cdot}4 = 3 : 2$.

3. *(a)* There are three urns having the following compositions of black and white balls.

Urn 1 : 7 white, 3 black balls
Urn 2 : 4 white, 6 black balls
Urn 3 : 2 white, 8 black balls.

One of these urns is chosen at random with probabilities 0·20, 0·60 and 0·20 respectively. From the chosen urn two balls are drawn at random without replacement. Calculate the probability that both these balls are white.

**Ans.** 8/45. **(Madurai Univ. B.Sc., 1991)**

*(b)* Bowl I contain 3 red chips and 7 blue chips, bowl II contain 6 red chips and 4 blue chips. A bowl is selected at random and then 1 chip is drawn from this bowl. *(i)* Compute the probability that this chip is red, *(ii)* Relative to the hypothesis that the chip is red, find the conditional probability that it is drawn from bowl II.

**[Delhi Univ. B.Sc. (Maths Hons.) 1987]**

*(c)* In a factory machines $A$ and $B$ are producing springs of the same type. Of this production, machines $A$ and $B$ produce 5% and 10% defective springs, respectively. Machines $A$ and $B$ produce 40% and 60% of the total output of the factory. One spring is selected at random and it is found to be defective. What is the possibility that this defective spring was produced by machine $A$?

**[Delhi Univ. M.A. (Econ.), 1986]**

*(d)* Urn $A$ contains 2 white, 1 black and 3 red balls, urn $B$ contains 3 white, 2 black and 4 red balls and urn $C$ contains 4 white, 3 black and 2 red balls. One urn is chosen at random and 2 balls are drawn. They happen to be red and black. What

is the probability that both balls came from urn '$B$' ?

[Madras U. B.Sc. April; 1989]

*(e)* Urn $X_1, X_2, X_3$, each contains 5 red and 3 white balls. Urns $Y_1, Y_2$, each contain 2 red and 4 white balls. An urn is selected at random and a ball is drawn. It is found to be red. Find the probability that the ball comes out of the urns of the first type. **[Bombay U. B.Sc., April 1992]**

*(f)* Two shipments of parts are received. The first shipment contains 1000 parts with 10% defectives and the second shipment contains 2000 parts with 5% defectives. One shipment is selected at random. Two parts are tested and found good. Find the probability (*a posterior*) that the tested parts were selected from the first shipment. **[Burdwan Univ. B.Sc. (Hons.), 1988]**

*(g)* There are three machines producing 10,000 ; 20,000 and 30,000 bullets per hour respectively. These machines are known to produce 5%, 4% and 2% defective bullets respectively. One bullet is taken at random from an hour's production of the three machines. What is the probability that it is defective? If the drawn bullet is defective, what is the probability that this was produced by the second machine? **[Delhi Univ. B.Sc. (Stat. Hons.), 1991]**

4. *(a)* Three urns are given each containing red and white chips as indicated.

Urn 1 : 6 red and 4 white.
Urn 2 : 2 red and 6 white.
Urn 3 : 1 red and 8 white.

*(i)* An urn is chosen at random and a ball is drawn from this urn. The ball is red. Find the probability that the urn chosen was urn I.

*(ii)* An urn is chosen at random and two balls are drawn without replacement from this urn. If both balls are red, find the probability that urn I was chosen. Under these conditions, what is the probability that urn III was chosen.

**Ans.** 108/173, 112/12, 0 **[Gauhati Univ. B.Sc., 1990]**

*(b)* There are ten urns of which each of three contains 1 white and 9 black balls, each of other three contains 9 white and 1 black ball, and of the remaining four each contains 5 white and 5 black balls. One of the urns is selected at random and a ball taken blindly from it turns out to be white. What is the probabililty that an urn containing 1 white and 9 black balls was selected? **(Agra Univ. B.Sc., 1991)**

**Hint :** $P(E_1) = \frac{3}{10}$, $P(E_2) = \frac{3}{10}$ and $P(E_3) = \frac{4}{10}$.

Let $A$ be the event of drawing a white ball.

$$P(A) = \frac{3}{10} \times \frac{1}{10} + \frac{3}{10} \times \frac{9}{10} + \frac{4}{10} \times \frac{5}{10} = \frac{1}{2}$$

$$P(A \mid E_1) = \frac{1}{10} \text{ and } P(E_1 \mid A) = \frac{3}{50}$$

*(c)* It is known that an urn containing altogether 10 balls was filled in the following manner: A coin was tossed 10 times, and according as it showed heads or tails, one white or one black ball was put into the urn. Balls are drawn from this

urn one at a time, 10 times in succession (with replacement) and every one turns out to be white. Find the chance that the urn contains nothing but white balls.

**Ans.** 0·0702.

5. (*a*) From a vessel containing 3 white and 5 black balls, 4 balls are transferred into an empty vessel. From this vessel a ball is drawn and is found to be white. What is the probability that out of four balls transferred, 3 are white and 1 black. **[Delhi Uni. B.Sc. (Stat. Hons.), 1985]**

**Hint.** Let the five mutually exclusive events for the four balls transferred be $E_0, E_1, E_2, E_3$, and $E_4$, where $E_i$ denotes the event that $i$ white balls are transferred and let $A$ be the event of drawing a white ball from the new vessel.

Then $$P(E_0) = \frac{^5C_4}{^8C_4},\ P(E_1) = \frac{^3C_1 \times {^5C_3}}{^8C_4},\ P(E_2) = \frac{^3C_2 \times {^5C_2}}{^8C_4}$$

$$P(E_3) = \frac{^3C_3 \times {^5C_1}}{^8C_4} \text{ and } P(E_4) = 0$$

Also $P(A \mid E_0) = 0,\ P(A \mid E_1) = \frac{1}{4},\ P(A \mid E_2) = \frac{2}{4},\ (A \mid E_3) = \frac{3}{4}$,

and $P(A \mid E_4) = 1$. Hence $P(E_3 \mid A) = \frac{1}{7}$.

(*b*) The contents of the urns 1 and 2 are as follows :

Urn 1 : 4 white and 5 black balls.

Urn 2 : 3 white and 6 black balls.

One urn is chosen at random and a ball is drawn and its colour noted and replaced back to the urn. Again a ball is drawn from the same urn, colour noted and replaced. The process is repeated 4 times and as a result one ball of white colour and three balls of black colour are obtained. What is the probability that the urn chosen was the urn 1 ? **(Poona Univ. B.E., 1989)**

**Hint.** $P(E_1) = P(E_2) = 1/2$,

$P(A \mid E_1) = 4/9, \qquad 1 - P(A \mid E_1) = 5/9$

$P(A \mid E_2) = 1/3, \qquad 1 - P(A \mid E_2) = 2/3$

The probability that the urn chosen was the urn 1

$$= \frac{\frac{1}{2} \cdot \frac{4}{9} \left(\frac{5}{9}\right)^3}{\frac{1}{2} \cdot \frac{4}{9} \cdot \left(\frac{5}{9}\right)^3 + \frac{1}{2} \cdot \frac{1}{3} \cdot \left(\frac{2}{3}\right)^3}$$

(*c*) There are five urns numbered 1 to 5. Each urn contains 10 balls. The $i$th urn has $i$ defective balls and $10 - i$ non-defective balls; $i = 1, 2, \ldots 5$. An urn is chosen at random and then a ball is selected at random from that urn. (*i*) What is the probability that a defective ball is selected ?

(*ii*) If the selected ball is defective, find the probability that it came from urn $i$, $(i = 1, 2, \ldots, 5)$. **[Delhi Univ. B.Sc. (Maths Hons.), 1987]**

**Hint.:** Define the following events :

$E_i$ : $i$th urn is selected at random.

$A$ : Defective ball is selected.

$P(E_i) = 1/5;\ i = 1, 2, \ldots, 5.$

$P(A \mid E_i) = P$ [Defective ball from $i$th urn] $= i/10,\ (i = 1, 2, \ldots, 5)$

$$P(E_i) \cdot P(A \mid E_i) = \frac{1}{5} \times \frac{i}{10} = \frac{i}{50}\ ,\quad (i = 1, 2, \ldots, 5).$$

(i) $$P(A) = \sum_{i=1}^{5} P(E_i)\, P(A \mid E_i) = \sum_{i=1}^{5} \left(\frac{i}{50}\right) = \frac{1+2+3+4+5}{50} = \frac{3}{10}$$

(ii) $$P(E_i \mid A) = \frac{P(E_i)\, P(A \mid E_i)}{\sum_i P(E_i)\, P(A \mid E_i)} = \frac{i/50}{3/10} = \frac{i}{15}\ ;\ i = 1, 2, \ldots, 5.$$

For example, the probability that the defective ball came from 5th urn $= (5/15) = 1/3$.

6. *(a)* A bag contains six balls of different colours and a ball is drawn from it at random. $A$ speaks truth thrice out of 4 times and $B$ speaks truth 7 times out of 10 times. If both $A$ and $B$ say that a red ball was drawn, find the probability of their joint statement being true.

**[Delhi Univ. B.Sc. (Stat. Hons.),1987; Kerala Univ. B.Sc.1988]**

*(b)* $A$ and $B$ are two very weak students of Statistics and their chances of solving a problem correctly are 1/8 and 1/12 respectively. If the probability of their making a common mistake is 1/1001 and they obtain the same answer, find the chance that their answer is correct. **[Poona Univ. B.Sc., 1989]**

**Ans.** Reqd. Probability $= \dfrac{1/8 \times 1/12}{1/8 \times 1/12 + (1 - 1/8) \cdot (1 - 1/12) \cdot 1/1001} = \dfrac{13}{14}$

7. *(a)* Three boxes, practically indistinguishable in appearance, have two drawers each. Box I contains a gold coin in one and a silver coin in the other drawer, box II contains a gold coin in each drawer and box III contains a silver coin in each drawer. One box is chosen at random and one of its drawers is opened at random and a gold coin found. What is the probability that the other drawer contains a coin of silver? **(Gujarat Univ. B.Sc., 1992)**

**Ans.** 1/3, 1/3.

*(b)* Two cannons No. 1 and 2 fire at the same target. Cannon No. 1 gives on an average 9 shots in the time in which Cannon No. 2 fires 10 projectiles. But on an average 8 out of 10 projectiles from Cannon No. 1 and 7 out of 10 from Cannon No. 2 strike the target. In the course of shooting, the target is struck by one projectile. What is the probability of a projectile which has struck the target belonging to Cannon No. 2 ? **(Lucknow Univ. B.Sc., 1991)**

**Ans.** 0·493

*(c)* Suppose 5 men out of 100 and 25 women out of 10,000 are colour blind. A colour blind person is chosen at random. What is the probability of his being male? (Assume males and females to be in equal number.)

**Hint.** $E_1$ = Person is a male, $E_2$ = Person is a female.

$A$ = Person is colour blind.

Then $P(E_1) = P(E_2) = \frac{1}{2}$, $P(A \mid E_1) = 0{\cdot}05$, $P(A \mid E_2) = 0{\cdot}0025$.

Hence find $P(E_1 \mid A)$.

8. *(a)* Three machines $X, Y, Z$ with capacities proportional to 2:3:4 are producting bullets. The probabilities that the machines produce defective are 0·1, 0·2 and 0·1 respectively. A bullet is taken from a day's production and found to be defective What is the probability that it came from machine $X$?

**[Madras Univ. B.Sc., 1988]**

*(b)* In a factory 2 machines $M_1$ and $M_2$ are used for manufacturing screws which may be uniquely classified as good or bad. $M_1$ produces per day $n_1$ boxes of screws, of which on the average, $p_1\%$ are bad while the corresponding numbers for $M_2$ are $n_2$ and $p_2$. From the total production of both $M_1$ and $M_2$ for a certain day, a box is chosen at random, a screw taken out of it and it is found to be bad. Find the chance that the selected box is manufactured *(i)* by $M_1$, *(ii)* $M_2$.

**Ans.** *(i)* $n_1 p_1/(n_1 p_1 + n_2 p_2)$, *(ii)* $n_2 p_2/(n_1 p_1 + n_2 p_2)$.

9. *(a)* A man is equally likley to choose any one of three routes $A, B, C$ from his house to the railway station, and his choice of route is not influenced by the weather. If the weather is dry, the probabilities of missing the train by routes $A, B, C$ are respectively 1/20, 1/10, 1/5. He sets out on a dry day and misses the train. What is the probability that the route chosen was $C$?

On a wet day, the respective probabilities of missing the train by routes $A, B, C$ are 1/20, 1/5, 1/2 respectively. On the average, one day in four is wet. If he misses the train, what is the probability that the day was wet?

**[Allahabad Univ. B.Sc., 1991]**

*(b)* A doctor is to visit the patient and from past experience it is known that the probabilities that he will come by train, bus or scooter are respectively 3/10, 1/5, and 1/10, the probabililty that he will use some other means of transport being, therefore, 2/5. If he comes by train, the probability that he will be late is 1/4, if by bus 1/3 and if by scooter 1/12, if he uses some other means of transport it can be assumed that he will not be late. When he arrives he is late. What is the probability that *(i)* he comes by train *(ii)* he is not late?

**[Burdwan Univ. B.Sc. (Hons.), 1990]**

**Ans.** *(i)* 1/2, *(ii)* 9/34

10. State and prove Bayes rule and expalin why, in spite of its easy deductibility from the postulates of probability, it has been the subject of such extensive controversy.

In the chest X-ray tests, it is found that the probability of detection when a person has actually T.B. is 0·95 and probability of diagnosing incorrectly as having T.B. is 0·002. In a certain city 0·1% of the adult population is suspected to be suffering from T.B. If an adult is selected at random and is diagnosed as having

T.B. on the basis of the X-ray test, what is the probability of his actually having a T.B.? **(Nagpur Univ. B.E., 1991)**

**Ans.** 0·97

11. A certain transistor is manufactured at three factories at Barnsley, Bradford and Bristol. It is known that the Barnsley factory produces twice as many transistors as the Bradford one, which produces the same number as the Bristol one (during the same period). Experience also shows that 0·2% of the transistors produced at Barnsley and Bradford are faulty and so are 0·4% of those produced at Bristol.

A service engineer, while maintaining an electronic equipment, finds a defective transistor. What is the probability that the Bradford factory is to blame?

**(Bangalore Univ. B.E., Oct. 1992)**

12. The sample space consists of integers from 1 to $2n$ which are assigned probabilities proportional to their logarithms. Find the probabilities and show that the conditional probability of the integer 2, given that an even integer occurs, is

$$\frac{\log 2}{[n \log 2 + \log (n!)]}$$ **(Lucknow Univ. M.A., 1992)**

[**Hint.** Let $E_i$ : the event that the integer $2i$ is drawn, $(i = 1, 2, 3, ..., n)$.

$A$ : the event of drawing an even integer.

$$\Rightarrow \quad A = E_1 \cup E_2 \cup \ldots \cup E_n \quad \Rightarrow \quad P(A) = \sum_{i=1}^{n} P(E_i)$$

But $P(E_i) = k \log (2i)$ (Given)

$$\therefore \quad P(A) = k \sum_{i=1}^{n} \log (2i) = k \log \prod_{i=1}^{n} (2i) = k [n \log 2 + \log (n!)]$$

$$\therefore \quad P(E_i \mid A) = \frac{\log (2i)}{[n \log 2 + \log (n!)]}$$

13. In answering a question on a multiple choice test, an examinee either knows the answer (with probability $p$), or he guesses (with probability $1 - p$). Assume that the probability of answering a question correctly is unity for an examinee who knows the answer and $1/m$ for the examinee who guesses, where $m$ is the number of multiple choice alternatives. Supposing an examinee answers a question correctly, what is the probability that he really knows the answer?

**[Delhi Univ. M.C.A., 1990; M.Sc. (Stat.), 1989]**

**Hint.** Let $E_1$ = The examinee knows the answer,

$E_2$ = The examinee guesses the answer,

and $A$ = The examinee answers correctly.

Then $P(E_1) = p$, $P(E_2) = 1 - p$, $P(A \mid E_1) = 1$ and $P(A \mid E_2) = 1/m$

Now use Bayes theorem to prove

$$P(E_1 \mid A) = \frac{mp}{1 + (m-1)p}$$

14. Die $A$ has four red and two white faces whereas die $B$ has two red and four white faces. A biased coin is flipped once. If it falls heads, the game continues by

throwing die $A$, if it falls tails die $B$ is to be used.

*(i)* Show that the probability of getting a red face at any throw is 1/2.

*(ii)* If the first two throws resulted in red faces, what is the probability of red face at the 3rd throw?

*(iii)* If red face turns up at the first $n$ throws, what is the probability that die $A$ is being used?

**Ans.** *(ii)* 3/5 *(iii)* $\dfrac{2^n}{2^n+1}$

**15.** A manufacturing firm produces steel pipes in three plants with daily production volumes of 500, 1,000 and 2,000 units respectively. According to past experience it is known that the fraction of defective outputs produced by the three plants are respectively 0.005, 0.008 and 0.010. If a pipe is selected at random from a day's total production and found to be defective, from which plant does that pipe come?

**Ans.** Third plant.

**16.** A piece of mechanism consists of 11 components, 5 of type $A$, 3 of type $B$, 2 of type $C$ and 1 of type $D$. The probability that any particular component will function for a period of 24 hours from the commencement of operations without breaking down is independent of whether or not any other component breaks down during that period and can be obtained from the following table:

*Component type*:ABCD

*Probability*:0·60·70·30·2

*(i)* Calculate the probability that 2 components chosen at random from the 11 components will both function for a period of 24 hours from the commencement of operations without breaking down.

*(ii)* If at the end of 24 hours of operations neither of the 2 components chosen in *(i)* has broken down, what is the probability that they are both type $C$ components.

**Hint.**

$$\begin{aligned}(i)\ \text{Required probability} &= \frac{1}{{}^{11}C_2}\,[\,{}^5C_2\times(0\cdot6)^2+{}^3C_2\,(0\cdot7)^2+{}^2C_2\,(0\cdot3)^2\\ &\quad+{}^5C_1\times{}^3C_1\times0\cdot6\times0\cdot7+{}^5C_1\times{}^2C_1\times(0\cdot6)\times(0\cdot3)\\ &\quad+{}^5C_1\times{}^1C_1\times(0\cdot6)\times(0\cdot2)+{}^3C_1\times{}^2C_1\times0\cdot7\times0\cdot3\\ &\quad+{}^3C_1\times{}^1C_1\times0\cdot7\times0\cdot2+{}^2C_1\times{}^1C_1\times0\cdot3\times0\cdot2]\\ &=p\ \text{(Say).}\end{aligned}$$

*(ii)* Required probability (By Bayes theorem)

$$=\frac{{}^2C_2\times(0\cdot3)^2}{p}=\frac{0\cdot09}{p}$$

**4·10. Geometric probability.** In remark 3, § 4·3·1 it was pointed out that the classical definition of probability fails if the total number of outcomes of an experiment is infinite. Thus, for example, if we are interested in finding the probability that a point selected at random in a given region will lie in a specified part of it, the classical definition of probability is modified and extended to what is called *geometrical probability or probability in continuum*. In this case, the general expression for probability '$p$' is given by

$$p = \frac{\textit{Measure of specified part of the region}}{\textit{Measure of the whole region}}$$

where 'measure' refers to the length, area or volume of the region if we are dealing with one, two or three dimensional space respectively.

**Example 4·34.** *Two points are taken at random on the given straight line of length a. Prove that the probability of their distance exceeding a given length* $c$ ($< a$) is equal to $\left(1 - \frac{c}{a}\right)^2$.

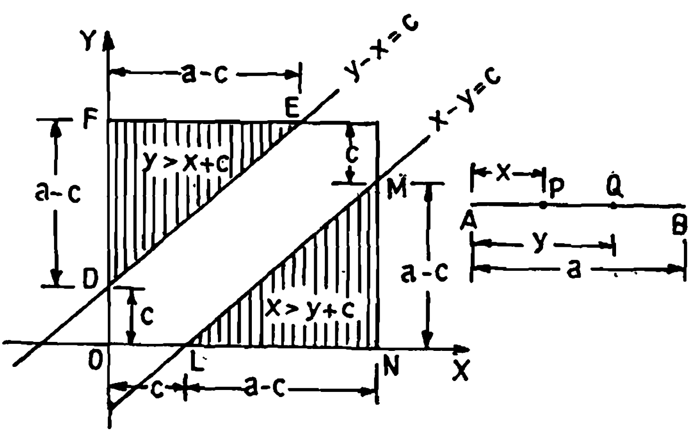
**[Burdwan Univ. B.Sc. (Hons.), 1992; Delhi Univ. M.A. (Econ.), 1987]**

**Solution.** Let $P$ and $Q$ be any two points taken at random on the given straight line $AB$ of length '$a$'. Let $AP = x$ and $AQ = y$,

$$(0 \le x \le a, 0 \le y \le a).$$

Then we want $P\{| x - y | > c\}$.

The probability can be easily calculated geometrically. Plotting the lines $x - y = c$ and $y - x = c$ along the co-ordinate axes, we get the following diagram:




Since $0 \le x \le a, 0 \le y \le a$, total area $= a \cdot a = a^2$.

Area favourable to the event $| x - y | > c$ is given by

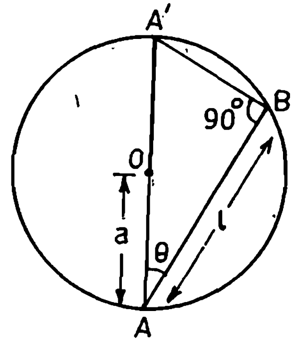
$$\Delta LMN + \Delta DEF = \frac{1}{2} LN \cdot MN + \frac{1}{2} EF \cdot DF$$

$$= \frac{1}{2}(a-c)^2 + \frac{1}{2}(a-c)^2 = (a-c)^2$$

$$P(| x - y | > c) = \frac{(a-c)^2}{a^2} = \left(1 - \frac{c}{a}\right)^2$$

**Example 4·35. (Bertrand's Problem).** *If a chord is taken at random in a circle, what is the chance that its length $l$ is not less than 'a', the radius of the circle?*

**Solution.** Let the chord $AB$ make an angle $\theta$ with the diameter $AOA'$ of the circle with centre $O$ and radius $OA=a$. Obviously $\theta$ lies betwseen $-\pi/2$ and $\pi/2$. Since all the positions of the chord $AB$ and consequently all the values of $\theta$ are equally likely, $\theta$ may be regarded as a random variable which is uniformly distributed *c.f.* § 8·1 over $(-\pi/2, \pi/2)$ with probability density function




$$f(\theta) = \frac{1}{\pi}\ ;\ -\pi/2 < \theta \leq \pi/2$$

$\angle ABA'$, being the angle in a semi-circle, is a right angle. From $\Delta\, ABA'$ we have

$$\frac{AB}{AA'} = \cos\theta$$

$$\Rightarrow \qquad l = 2a\cos\theta$$

The required probability '$p$' is given by

$$p = P(l \geq a) = P(2a\cos\theta \geq a)$$
$$= P(\cos\theta \geq 1/2) = P(|\theta| \leq \pi/3)$$
$$= \int_{-\pi/3}^{\pi/3} f(\theta)\, d\theta = \frac{1}{\pi}\int_{-\pi/3}^{\pi/3} d\theta = \frac{2}{3}$$

**Example 4·36.** *A rod of length 'a' is broken into three parts at random. What is the probability that a triangle can be formed from these parts?*

**Solution.** Let the lengths of the three parts of the rod be $x$, $y$ and $a-(x+y)$. Obviously, we have

$$x > 0\ ;\ y > 0 \text{ and } x + y < a \quad \Rightarrow \quad y < a - x \qquad \ldots(*)$$

In order that these three parts form the sides of a triangle, we should have

$$\left.\begin{array}{lll} x+y > a-(x+y) & \Rightarrow & y > \frac{a}{2} - x \\ \text{and} \quad x + a - (x+y) > y & \Rightarrow & y < \frac{a}{2} \\ y + a - (x+y) > x & \Rightarrow & y < \frac{a}{2} \end{array}\right\} \qquad \ldots(**)$$

since in a triangle, the sum of any two sides is greater than the third. Equivalently, (**) can be written as

$$\frac{a}{2} - x < y < \frac{a}{2} \quad \wedge \quad 0 < x < \frac{a}{2} \qquad \ldots(***)$$

Hence, on using (*) and (***), the required probability is given by

$$\frac{\int_0^{a/2}\int_{(a/2)-x}^{a/2} dy\,dx}{\int_0^{a}\int_0^{a-x} dy\,dx} = \frac{\int_0^{a/2}\left[\frac{a}{2}-\left(\frac{a}{2}-x\right)\right]dx}{\int_0^{a}(a-x)\,dx}$$

$$= \frac{\left|\frac{x^2}{2}\right|_0^{a/2}}{\left|\frac{-(a-x)^2}{2}\right|_0^{a}} = \frac{a^2/8}{a^2/2} = \frac{1}{4}$$

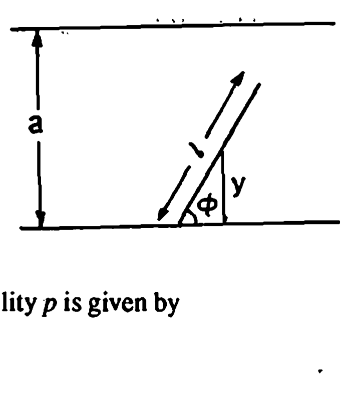
**Example 4·37. (Buffon's Needle Problem).** *A vertical board is ruled with horizontal parallel lines at constant distance 'a' apart. A needle of length $l$ $(< a)$ is thrown at random on the table. Find the probability that it will intersect one of the lines.*

**Solution.** Let $y$ denote the distance from the centre of the needle to the nearest parallel and $\phi$ be angle formed by the needle with this parallel. The quantities $y$ and $\phi$ fully determine the position of the needle. Obviously $y$ ranges from 0 to $a/2$ (since $l < a$) and $\phi$ from 0 to $\pi$.

Since the needle is dropped randomly, all possible values of $y$ and $\phi$ may be regarded as equally likely and consequently the joint probability density function $f(y, \phi)$ of $y$ and $\phi$ is given by the uniform distribution ( c.f. § 8.1 ) by

$$f(y, \phi) = k; \quad 0 \le \phi \le \pi, \quad 0 \le y \le a/2, \quad ...(*)$$

where $k$ is a constant.




The needle will intersect one of the lines if the distance of its centre from the line is less than $\frac{1}{2}\, l \sin\phi$, *i.e.*, the required event can be represented by the inequality $0 < y < \frac{1}{2}\, l \sin\phi$. Hence the required probability $p$ is given by

$$p = \frac{\int_0^{\pi}\int_0^{(l\sin\phi)/2} f(y, \phi)\, dy\, d\phi}{\int_0^{\pi}\int_0^{a/2} f(y, \phi)\, dy\, d\phi}$$

$$= \frac{\frac{l}{2}\int_0^{\pi} \sin\phi\, d\phi}{(a/2)\,.\,\pi}$$

$$= \frac{l}{a\pi}\left|-\cos\phi\right|_0^{\pi} = \frac{2l}{a\pi}$$

## EXERCISE 4 (e)

1. Two points are selected at random in a line $AC$ of length '$a$' so as to lie on the opposite sides of its mid-point $O$. Find the probability that the distance between them is less than $a/3$.

2. *(a)* Two points are selected at random on a line of length $a$. What is the probability that none of three sections in which the line is thus divided is less than $a/4$ ?

Ans. $1/16$.

*(b)* A rectilinear segment $AB$ is divided by a point $C$ into two parts $AC=a$, $CB=b$. Points $X$ and $Y$ are taken at random on $AC$ and $CB$ respectively. What is the probability that $AX, XY$ and $BY$ can form a triangle?

*(c)* $ABG$ is a straight line such that $AB$ is 6 inches and $BG$ is 5 inches. A point $Y$ is chosen at random on the $BG$ part of the line. If $C$ lies between $B$ and $G$ in such a way that $AC=t$ inches, find

*(i)* the probability that $Y$ will lie in $BC$.

*(ii)* the probability that $Y$ will lie in $CG$.

What can you say about the sum of these probabilities?

*(d)* The sides of a rectangle are taken at random each less than $a$ and all lengths are equally likely. Find the chance that the diagonal is less than $a$.

3. *(a)* Three points are taken at random on the circumference of a circle. Find the chance that they lie on the same semi- circle.

*(b)* A chord is drawn at random in a given circle. What is the probability that it is greater than the side of an equilateral triangle inscribed in that circle?

*(c)* Show that the probability of choosing two points randomly from a line segment of length 2 inches and their being at a distance of at least 1 inch from each other is 1/4. [Delhi Univ. M.A. (Econ.), 1985]

4. A point is selected at random inside a circle. Find the probability that the point is closer to the centre of the circle than to its circumference.

5. One takes at random two points $P$ and $Q$ on a segment $AB$ of length $a$

*(i)* What is the probability for the distance $PQ$ being less than $b$ ( $<a$ )?

*(ii)* Find the chance that the distance between them is greater than a given length $b$.

6. Two persons $A$ and $B$, make an appointment to meet on a certain day at a certain place, but without fixing the time further than that it is to be between 2 p.m. and 3 p.m and that each is to wait not longer than ten minutes for the other. Assuming that each is independently equally likely to arrive at any time during the hour, find the probability that they meet.

Third person $C$, is to be at the same place from 2·10 p.m. until 2·40 p.m. on the same day. Find the probabilities of $C$ being present when $A$ and $B$ are there together *(i)* When $A$ and $B$ remain after they meet, *(ii)* When $A$ and $B$ leave as soon as they meet.

**Hint.** Denote the times of arrival of $A$ by $x$ and of $B$ by $y$. For the meeting to take place it is necessary and sufficient that

$$| x-y | < 10$$

We depict $x$ and $y$ as Cartesian coordinates in the plane; for the scale unit we take one minute. All possible outcomes can be described as points of a aquare with side 60. We shall finally get [*c.f.* Example 4·34, with $a = 60$, $c = 10$]

$$P[|x-y|<10] = 1-(5/6)^2 = 11/36$$

**7.** The outcome of an experiment are represented by points in the square bounded by $x = 0$, $x = 2$ and $y = 2$ in the $xy$-plane. If the probability is distributed uniformly, determine the probability that $x^2 + y^2 > 1$

**Hint.**

Required probability $P(E) = \int_E \frac{1}{4}\, dx\, dy = 1 - \int_{E'} \frac{1}{4}\, dx\, dy$ .

where $E$ is the region for which $x^2 + y^2 > 1$ and $E'$ is the region for which $x^2 + y^2 \leq 1$.

$$\therefore \quad 4P(E) = 4 - \int_0^1 \int_0^1 dx\, dy = 3 \quad \Rightarrow \quad P(E) = \frac{3}{4}$$

**8.** A floor is paved with tiles, each tile being a parallelogram such that the distance between pairs of opposite sides are $a$ and $b$ respectively, the length of the diagonal being $l$. A stick of length $c$ falls on the floor parallel to the diagonal. Show that the probability that it will lie entirely on one tile is

$$\left(1-\frac{c}{l}\right)^2$$

If a circle of diameter $d$ is thrown on the floor, show that the probability that it will lie on one tile is

$$\left(1-\frac{d}{a}\right)\left(1-\frac{d}{b}\right)$$

**9.** Circular discs of radius $r$ are thrown at random on to a plane circular table of radius $R$ which is surrounded by a border of uniform width $r$ lying in the same plane as the table. If the discs are thrown independently and at random, and $N$ stay on the table, show that the probability that a fixed point on the table but not on the border, will be covered is

$$1-\left(1-\frac{r^2}{(R+r)^2}\right)^N$$

## SOME MISCELLANEOUS EXAMPLES

**Example 4·38.** *A die is loaded in such a manner that for n=1, 2, 3, 4, 5, 6 the probability of the face marked n, landing on top when the die is rolled is proportional to n. Find the probability that an odd number will appear on tossing the die.*

**[Madras Univ. B.Sc. (Stat. Main), 1987]**

**Solution.** Here we are given

$P(n) \propto n$ or $P(n) = kn$, where $k$ is the constant of proportionality.

Also $P(1) + P(2) + \ldots P(6) = 1 \Rightarrow k(1+2+3+4+5+6) = 1$ or $k = 1/21$

Required Probability $= P(1) + P(3) + P(5) = \dfrac{1+3+5}{21} = \dfrac{3}{7}$

**Example 4·39.** *In terms of probability :*

$$p_1 = P(A),\ p_2 = P(B),\ p_3 = P(A \cap B),\ (p_1, p_2, p_3 > 0)$$

Express the following in terms of $p_1, p_2, p_3$

(a) $P(\overline{A \cup B})$, (b) $P(\bar{A} \cup \bar{B})$, (c) $P(\bar{A} \cap B)$, (d) $P(\bar{A} \cup B)$, (e) $P(\bar{A} \cap \bar{B})$

(f) $P(A \cap \bar{B})$, (g) $P(A \mid B)$, (h) $P(B \mid \bar{A})$, (i) $P[\bar{A} \cap (A \cup B)]$

**Solution.**

(a) $P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - [P(A) + P(B) - P(AB)]$

$= 1 - p_1 - p_2 + p_3.$

(b) $P(\bar{A} \cup \bar{B}) = P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - p_3$

(c) $P(\bar{A} \cap B) = P(B - AB) = P(B) - P(A \cap B) = p_2 - p_3$

(d) $P(\bar{A} \cup B) = P(\bar{A}) + P(B) - P(\bar{A} \cap B) = 1 - p_1 + p_2 - (p_2 - p_3)$

$= 1 - p_1 + p_3$

(e) $P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - p_1 - p_2 + p_3.$ [Part (a)]

(f) $P(A \cap \bar{B}) = P(A - A \cap B) = P(A) - P(A \cap B) = p_1 - p_3$

(g) $P(A \mid B) = P(A \cap B)/P(B) = p_3/p_2$

(h) $P(B \mid \bar{A}) = P(\bar{A} \cap B)/P(\bar{A}) = (p_2 - p_3)/(1 - p_1)$

(i) $P[\bar{A} \cap (A \cup B)] = P[(\bar{A} \cap A) \cup (\bar{A} \cap B)]$

$= P(\bar{A} \cap B) = p_2 - p_3$ $[\because A \cap \bar{A} = \phi]$

**Example 4·40.** *Let $P(A) = p$, $P(A \mid B) = q$, $P(B \mid A) = r$. Find relations between the numbers $p, q, r$ for the following cases :*

(a) *Events A and B are mutually exclusive.*
(b) *A and B are mutually exclusive and collectively exhaustive.*
(c) *A is a subevent of B; B is a subevent of A.*
(d) *$\bar{A}$ and $\bar{B}$ are mutually exclusive.*

**[Delhi Univ. B.Sc. (Maths Hons.) 1985]**

**Solution.** From given data : $P(A) = p$, $P(A \cap B) = P(A)\, P(B \mid A) = rp$

$$\therefore \quad P(B) = \frac{P(A \cap B)}{P(A \mid B)} = \frac{rp}{q}$$

(a) $P(A \cap B) = 0 \quad \Rightarrow \quad rp = 0.$

(b) $P(A \cap B) = 0$ and $P(A) + P(B) = 1$

$\Rightarrow p(q + r) = q;\ rp = 0 \quad \Rightarrow \quad pq = q \quad \Rightarrow \quad p = 1 \vee q = 0.$

(c) $A \subseteq B \Rightarrow A \cap B = A$ or $P(A \cap B) = P(A) \Rightarrow rp = p \Rightarrow r = 1 \vee p = 0.$

$B \subseteq A \Rightarrow A \cap B = B$ or $P(A \cap B) = P(B)$

$\Rightarrow \quad rp = (rp/q)$ or $rp(q - 1) = 0 \quad \Rightarrow \quad q = 1$

(d) $P(\bar{A} \cap \bar{B}) = 1 - P(A \cup B) \quad \Rightarrow \quad 0 = 1 - [P(A) + P(B) - P(A \cap B)]$

So $P(A)+P(B)=1+P(A\cap B) \Rightarrow p\,[1+(r/q)]=1+rp$

$\therefore \quad p\,(q+r)=q\,(1+pr).$

**Example 4·41.** *(a) Twelve balls are distributed at random among three boxes. What is the probability that the first box will contain 3 balls?*

*(b) If n biscuits be distributed among N persons, find the chance that a particular person receives r ( < n ) biscuits.* **[Marathwada Univ. B.Sc. 1992]**

**Solution.** *(a)* Since each ball can go to any one of the three boxes, there are 3 ways in which a ball can go to any one of the three boxes. Hence there are $3^{12}$ ways in which 12 balls can be placed in the three boxes.

Number of ways in which 3 balls out of 12 can go to the first box is $^{12}C_3$. Now the remaining 9 balls are to be placed in 2 boxes and this can be done in $2^9$ ways. Hence the total number of favourable cases $= {}^{12}C_3 \times 2^9$.

$$\therefore \quad \text{Required probability} = \frac{^{12}C_3 \times 2^9}{3^{12}}$$

*(b)* Take any one biscuit. This can be given to any one of the $N$ beggars so that there are $N$ ways of distributing any one biscuit. Hence the total number of ways in which $n$ biscuit can be distributed at random among $N$ beggars

$$= N \,.\, N \ldots N \,(\,n \text{ times}\,) = N^n.$$

'$r$' biscuits can be given to any particular beggar in $^nC_r$ ways. Now we are left with $(n-r)$ biscuits which are to be distributed among the remaining $(N-1)$ beggars and this can be done in $(N-1)^{n-r}$ ways.

$\therefore \quad$ Number of favourable cases $= {}^nC_r \,.\, (N-1)^{n-r}$

$$\text{Hence, required probability} = \frac{^nC_r\,(N-1)^{n-r}}{N^n}$$

**Example 4·42.** *A car is parked among N cars in a row, not at either end. On his return the owner finds that exactly r of the N places are still occupied. What is the probability that both neighbouring places are empty?*

**Solution.** Since the owner finds on return that exactly $r$ of the $N$ places (including owner's car) are occupied, the exhaustive number of cases for such an arrangement is $^{N-1}C_{r-1}$ [since the remaining $r-1$ cars are to be parked in the remaining $N-1$ places and this can be done in $^{N-1}C_{r-1}$ ways].

Let $A$ denote the event that both the neighbouring places to owner's car are empty. This requires the remaining $(r-1)$ cars to be parked in the remaining $N-3$ places and hence the number of cases favourable to $A$ is $^{N-3}C_{r-1}$. Hence

$$P(A) = \frac{^{N-3}C_{r-1}}{^{N-1}C_{r-1}} = \frac{(N-r)(N-r-1)}{(N-1)(N-2)}$$

**Example 4·43.** *What is the probability that at least two out of n people have the same birthday? Assume 365 days in a year and that all days are equally likely.*

**Solution.** Since the birthday of any person can fall on any one of the 365 days, the exhaustive number of cases for the birthdays of $n$ persons is $365^n$.

If the birthdays of all the $n$ persons fall on different days, then the number of favourable cases is

$$365\,(365-1)\,(365-2)\,....\,[365-(n-1)],$$

because in this case the birthday of the first person can fall on any one of 365 days, the birthday of the second person can fall on any one of the remaining 364 days and so on.

Hence the probability $(p)$ that birthdays of all the $n$ persons are different is given by :

$$p = \frac{365\,(365-1)\,(365-2)\ldots[365-(n-1)]}{365^n}$$

$$= \left(1-\frac{1}{365}\right)\left(1-\frac{2}{365}\right)\left(1-\frac{3}{365}\right)\cdots\left(1-\frac{n-1}{365}\right)$$

Hence the required probability that at least two persons have the same birthday is

$$1-p = 1-\left(1-\frac{1}{365}\right)\left(1-\frac{2}{365}\right)\left(1-\frac{3}{365}\right)\cdots\left(1-\frac{n-1}{365}\right)$$

**Example 4·44.** *A five-figure number is formed by the digits 0, 1, 2, 3, 4 (without repetition). Find the probability that the number formed is divisible by 4.*

**[Delhi Univ. B.Sc. (Stat. Hons.), 1990]**

**Solution.** The total number of ways in which the five digits 0, 1, 2, 3, 4 can be arranged among themselves is 5!. Out of these, the number of arrangements which begin with 0 (and, therefore, will give only 4-digited numbers) is 4!. Hence the total number of five digited numbers that can be formed from the digits 0, 1, 2, 3, 4 is

$$5\,!-4\,!=120-24=96$$

The number formed will be divisible by 4 if the number formed by the two digits on extreme right (*i.e.*, the digits in the unit and tens places) is divisible by 4. Such numbers are :

$$04\ ,\ 12\ ,\ 20\ ,\ 24\ ,\ 32\ ,\ \text{and}\ 40$$

If the numbers end in 04, the remaining three digits, *viz.*, 1, 2 and 3 can be arranged among themselves in 3 ! ways. Similarly, the number of arrangements of the numbers ending with 20 and 40 is 3 ! in each case.

If the numbers end with 12, the remaining three digits 0, 3, 4 can be arranged in 3 ! ways. Out of these we shall reject those numbers which start with 0 (*i.e.*, have 0 as the first digit). There are $(3-1)\,!=2\,!$ such cases. Hence, the number of five digited numbers ending with 12 is

$$3\,!-2\,!=6-2=4$$

Similarly the number of 5 digited numbers ending with 24 and 32 each is 4. Hence the total number of favourable cases is

$$3 \times 3\,! + 3 \times 4 = 18 + 12 = 30$$

Hence required probability $= \frac{30}{96} = \frac{5}{16}$

**Example 4·45.** *(Huyghen's problem). A and B throw alternately with a pair of ordinary dice. A wins if he throws 6 before B throws 7, and B wins if he throws 7 before A throws 6. If A begins, show that his chance of winning is 30 / 61*

**[Delhi Univ. B.Sc. (Stat. Hons.), 1991; Delhi Univ. B.Sc.,1987]**

**Solution.** Let $E_1$ denote the event of $A$'s throwing '6' and $E_2$ the event of $B$'s throwing '7 with a pair of dice. Then $\bar{E}_1$ and $\bar{E}_2$ are the complementary events.

'6' can be obtained with two dice in the following ways:

(1, 5 ), (5, 1 ), (2, 4), (4, 2), (3, 3), *i.e.*, in 5 distinct ways.

$$\therefore \qquad P(E_1) = \frac{5}{36} \quad \text{and} \quad P(\bar{E}_1) = 1 - \frac{5}{36} = \frac{31}{36}$$

'7' can be obtained with two dice as follows:

(1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3), *i.e.*, in 6 distinct ways.

$$\therefore \qquad P(E_2) = \frac{6}{36} = \frac{1}{6} \quad \text{and} \quad P(\bar{E}_2) = 1 - \frac{1}{6} = \frac{5}{6}$$

If $A$ starts the game, he will win in the following mutually exclusive ways:

*(i)* $E_1$ happens *(ii)* $\bar{E}_1 \cap \bar{E}_2 \cap E_1$ happens

*(iii)* $\bar{E}_1 \cap \bar{E}_2 \cap \bar{E}_1 \cap \bar{E}_2 \cap E_1$ happens, and so on.

Hence by addition theorem of probability, the required probability of $A$'s winning, (say), $P(A)$ is given by

$$P(A) = P(i) + P(ii) + P(iii) + \ldots$$
$$= P(E_1) + P(\bar{E}_1 \cap \bar{E}_2 \cap E_1) + P(\bar{E}_1 \cap \bar{E}_2 \cap \bar{E}_1 \cap \bar{E}_2 \cap E_1) + \ldots.$$
$$= P(E_1) + P(\bar{E}_1)\, P(\bar{E}_2)\, P(E_1) + P(\bar{E}_1)\, P(\bar{E}_2)\, P(\bar{E}_1)\, P(\bar{E}_2)\, P(E_1) + \ldots$$

(By compound probability theorem)

$$= \frac{5}{36} + \frac{31}{36} \times \frac{5}{6} \times \frac{5}{36} + \frac{31}{36} \times \frac{5}{6} \times \frac{31}{36} \times \frac{5}{6} \times \frac{5}{36} + \ldots$$
$$= \frac{5/36}{1 - \frac{31}{36} \times \frac{5}{6}} = \frac{30}{61}$$

**Example 4·46.** *A player tosses a coin and is to score one point for every head and two points for every tail turned up. He is to play on until his score reaches or passes n. If $p_n$ is the chance of attaining exactly n score, show that*

$$p_n = \frac{1}{2}\,[\,p_{n-1} + p_{n-2}\,],$$

*and hence find the value of $p_n$.* **[Delhi Univ. B.Sc. (Stat. Hons.),1992]**

**Solution.** The score $n$ can be reached in the following two mutually exclusive ways:

*(i)* By throwing a tail when score is $(n-2)$, and

*(ii)* By throwing a head when score is $(n-1)$.

Hence by addition theorem of probability, we get

$$p_n = P(i) + P(ii) = \frac{1}{2} \cdot p_{n-2} + \frac{1}{2} \cdot p_{n-1} = \frac{1}{2}(p_{n-1} + p_{n-2}) \qquad \ldots(*)$$

To find $p_n$ explicitly, $(*)$ may be re-written as

$$\begin{aligned} p_n + \frac{1}{2} p_{n-1} &= p_{n-1} + \frac{1}{2} p_{n-2} \\ &= p_{n-2} + \frac{1}{2} p_{n-3} \\ &\ldots \quad \ldots \\ &\ldots \quad \ldots \\ &= p_2 + \frac{1}{2} p_1 \end{aligned} \qquad \ldots(**)$$

Since the score 2 can be obtained as

*(i)* Head in first throw and head in 2nd throw,

*(ii)* Tail in the first throw, we have

$$p_2 = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} = \frac{1}{4} + \frac{1}{2} = \frac{3}{4} \text{ and obviously } p_1 = \frac{1}{2}$$

Hence, from $(**)$, we get

$$p_n + \frac{1}{2} p_{n-1} = \frac{3}{4} + \frac{1}{2} \cdot \frac{1}{2} = 1 = \frac{2}{3} + \frac{1}{3} = \frac{2}{3} + \frac{1}{2} \cdot \frac{2}{3}$$

$$\left.\begin{aligned} p_n - \tfrac{2}{3} &= \left(-\tfrac{1}{2}\right)\left(p_{n-1} - \tfrac{2}{3}\right) \\ p_{n-1} - \tfrac{2}{3} &= \left(-\tfrac{1}{2}\right)\left(p_{n-2} - \tfrac{2}{3}\right) \\ &\vdots \qquad\qquad \vdots \\ p_2 - \tfrac{2}{3} &= \left(-\tfrac{1}{2}\right)\left(p_1 - \tfrac{2}{3}\right) \end{aligned}\right\}$$

Multiplying all the above equations, we get

$$\begin{aligned} p_n - \tfrac{2}{3} &= \left(-\tfrac{1}{2}\right)^{n-1}\left(p_1 - \tfrac{2}{3}\right) \\ &= \left(-\tfrac{1}{2}\right)^{n-1}\left(\tfrac{1}{2} - \tfrac{2}{3}\right) = (-1)^n \cdot \frac{1}{2^n} \cdot \frac{1}{3} \end{aligned}$$

$$\Rightarrow \qquad \begin{aligned} p_n &= \frac{2}{3} + (-1)^n \frac{1}{2^n} \cdot \frac{1}{3} \\ &= \frac{1}{3}\left[2 + (-1)^n \frac{1}{2^n}\right] \end{aligned}$$

**Example 4·47.** *A coin is tossed $(m+n)$ times, $(mn)$. Show that the probability of at least m consecutive heads is* $\dfrac{n+2}{2^{m+1}}$.

**[Kurukshetra Univ. M.Sc. 1990; Calcutta Univ. B.Sc.(Hons.), 1986]**

**Solution.** Since $m > n$, only one sequence of $m$ consecutive heads is possible. This sequence may start either with the first toss or second toss or third toss, and so on, the last one will be starting with $(n+1)th$ toss.

Let $E_i$ denote the event that the sequence of $m$ consecutive heads starts with $i$th toss. Then the required probability is

$$P(E_1) + P(E_2) + \ldots + P(E_{n+1}) \qquad \ldots(*)$$

Now $P(E_1) = P$ [Consecutive heads in first $m$ tosses and head or tail in the rest]

$$= \left(\frac{1}{2}\right)^m$$

$P(E_2) = P$ [Tail in the first toss, followed by $m$ consecutive heads and head or tail in the next]

$$= \frac{1}{2}\left(\frac{1}{2}\right)^m = \frac{1}{2^{m+1}}$$

In general,

$P(E_r) = P$ [tail in the $(r-1)$th trial followed by $m$ consecutive heads and head or tail in the next ]

$$= \frac{1}{2}\left(\frac{1}{2}\right)^m = \frac{1}{2^{m+1}}, \quad \forall \quad r = 2, 3, \ldots, n+1.$$

Substituting in (*),

$$\text{Required probability} = \frac{1}{2^m} + \frac{n}{2^{m+1}} = \frac{2+n}{2^{m+1}}$$

**Example 4·48.** *Cards are dealt one by one from a well-shuffled pack until an ace appears. Show that the probability that exactly n cards are dealt before the first ace appears is*

$$\frac{4(51-n)(50-n)(49-n)}{52.51.50.49}$$

**[Delhi Univ. B.Sc. 1992]**

**Solution.** Let $E_i$ denote the event that an ace appears when the $i$th card is dealt. Then the required probability '$p$' is given by

$p = P$ [Exactly $n$ cards are dealt before the first ace appears]

$= P$ [The first ace appears at the $(n+1)$th dealing]

$$= P(\bar{E}_1 \cap \bar{E}_2 \cap \bar{E}_3 \cap \ldots \cap \bar{E}_{n-1} \cap \bar{E}_n \cap E_{n+1})$$

$$= P(\bar{E}_1)\, P(\bar{E}_2 \mid \bar{E}_1)\, P(\bar{E}_3 \mid \bar{E}_1 \cap \bar{E}_2) \ldots$$

$$\times P(\bar{E}_n \mid \bar{E}_1 \cap \bar{E}_2 \cap \ldots \cap \bar{E}_{n-1}) \times P(E_{n+1} \mid \bar{E}_1 \cap \bar{E}_2 \cap \ldots \cap \bar{E}_n) \qquad \ldots(*)$$

Now

$$P(E_1) = \frac{4}{52} \quad \Rightarrow \quad P(\bar{E}_1) = \frac{48}{52}$$

$$P(E_2 \mid \bar{E}_1) = \frac{4}{51} \quad \Rightarrow \quad P(\bar{E}_2 \mid \bar{E}_1) = \frac{47}{51}$$

$$P(E_3 \mid \bar{E}_1 \cap \bar{E}_2) = \frac{4}{50} \quad \Rightarrow \quad P(\bar{E}_3 \mid \bar{E}_1 \cap \bar{E}_2) = \frac{46}{50}$$

$$\vdots \qquad\qquad \vdots$$

$$P(E_{n-1} \mid \bar{E}_1 \cap \bar{E}_2 \cap \ldots \cap \bar{E}_{n-2}) = \frac{4}{52-(n-2)}$$

$$\therefore \quad P(\bar{E}_{n-1} \mid \bar{E}_1 \cap \bar{E}_2 \cap \ldots \cap \bar{E}_{n-2}) = \frac{50-n}{52-(n-2)}$$

$$P(E_n \mid \bar{E}_1 \cap \bar{E}_2 \cap \ldots \cap \bar{E}_{n-1}) = \frac{4}{52-(n-1)}$$

$$\therefore \quad P(\bar{E}_n \mid \bar{E}_1 \cap \bar{E}_2 \cap \ldots \cap \bar{E}_{n-1}) = \frac{49-n}{52-(n-1)}$$

and $\quad P(E_{n+1} \mid \bar{E}_1 \cap \bar{E}_2 \cap \ldots \cap \bar{E}_n) = \dfrac{4}{52-n}$

Hence, from (*) we get

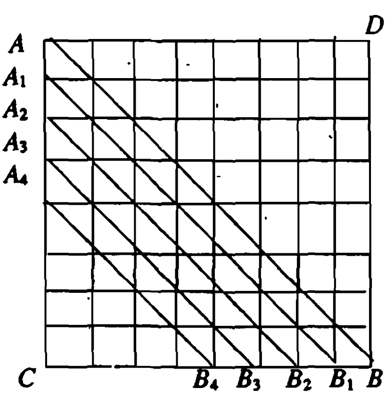
$$p = \left[\frac{48}{52} \times \frac{47}{51} \times \frac{46}{50} \times \frac{45}{49} \times \frac{44}{48} \times \frac{43}{47} \times \ldots \times \frac{52-n}{52-(n-4)} \right.$$
$$\left. \times \frac{51-n}{52-(n-3)} \times \frac{50-n}{52-(n-2)} \times \frac{49-n}{52-(n-1)} \times \frac{4}{52-n}\right]$$

$$= \frac{(51-n)(50-n)(49-n)\,4}{52 \times 51 \times 50 \times 49}$$

**Example 4·49.** *If four squares are chosen at random on a chess-board, find the chance that they should be in a diagonal line.*

**[Delhi Univ. B.Sc. (Stat. Hons.), 1988]**

**Solution.** In a chess-board there are $8 \times 8 = 64$ squares as shown in the following diagram.




Let us consider the number of ways in which the 4 squares selected at random are in a diagonal line parallel to $AB$. Consider the $\Delta\, ABC$. Number of ways in which 4 selected squares are along the lines $A_4 B_4$, $A_3 B_3$, $A_2 B_2$, $A_1 B_1$ and $AB$ are ${}^4C_4$, ${}^5C_4$, ${}^6C_4$, ${}^7C_4$ and ${}^8C_4$ respectively.

Similarly, in $\Delta\, ABD$ there are an equal number of ways of selecting 4 squares in a diagonal line parallel to $AB$.

Hence, total number of ways in which the 4 selected squares are in a diagonal line parallel to $AB$ are $2\,({}^4C_4 + {}^5C_4 + {}^6C_4 + {}^7C_4) + {}^8C_4$.

Since there is an equal number of ways in which 4 selected squares are in a diagonal line parallel to $CD$, the required number of favourable cases is given by

$$2\,[\,2({}^4C_4 + {}^5C_4 + {}^6C_4 + {}^7C_4) + {}^8C_4\,]$$

Since 4 squares can be selected out of 64 in ${}^{64}C_4$ ways, the required probability is

$$= \frac{2\,[\,2({}^4C_4 + {}^5C_4 + {}^6C_4 + {}^7C_4) + {}^8C_4]}{{}^{64}C_4}$$

$$= \frac{[4\,(\,1+5+15+35\,) + 140\,] \times 4\,!}{64 \times 63 \times 62 \times 61} = \frac{91}{158844}$$

**Example 4·50.** *An urn contains four tickets marked with numbers 112, 121, 211, 222 and one ticket is drawn at random. Let $A_i$, $(i=1,2,3)$ be the event that ith digit of the number of the ticket drawn is 1. Discuss the independence of the events $A_1, A_2$ and $A_3$.* **[Delhi Univ. B.Sc.(Stat. Hons.),1987; Poona Univ. B.Sc.,1986]**

**Solution.** We have

$$P(A_1) = \frac{2}{4} = \frac{1}{2} = P(A_2) = P(A_3)$$

$A_1 \cap A_2$ is the event that the first two digits in the number which the selected ticket bears are each equal to unity and the only favourable case is ticket with number 112.

$$\therefore \qquad P(A_1 \cap A_2) = \frac{1}{4} = \frac{1}{2}\cdot\frac{1}{2}$$

$$= P(A_1)\,P(A_2)$$

Similarly,

$$P(A_2 \cap A_3) = \frac{1}{4} = P(A_2)\,P(A_3)$$

and $$P(A_3 \cap A_1) = \frac{1}{4} = P(A_3)\,P(A_1)$$

Thus we conclude that the events $A_1, A_2$ and $A_3$ are pairwise independent.

Now $P(A_1 \cap A_3 \cap A_3) = P$ {all the three digits in the number are 1's}

$$= P(\phi)$$

$$= 0 \neq P(A_1)\,P(A_2)\,P(A_3)$$

Hence $A_1, A_2$ and $A_3$ though pairwise independent are not mutually independent.

**Example 4·51.** *Two fair dice are thrown independently. Three events A, B and C are defined as follows:*

*A : Odd face with first dice*

*B : Odd face with second dice*

*C : Sum of points on two dice is odd.*

*Are the events A, B and C mutually independent?*

**[Delhi Univ. B.Sc. (Stat. Hons.) 1983; M.S. Baroda Univ. B.Sc.1987]**

**Solution.** Since each of the two dice can show any one of the six faces 1, 2, 3, 4, 5, 6, we get :

$$P(A) = \frac{3 \times 6}{36} = \frac{1}{2} \qquad [\because A = \{1, 3, 5\} \times \{1, 2, 3, 4, 5, 6\}]$$

$$P(B) = \frac{3 \times 6}{36} = \frac{1}{2} \qquad [\because B = \{1, 2, 3, 4, 5, 6\} \times \{1, 3, 5\}]$$

The sum of points on two dice will be odd if one shows odd number and the other shows even number. Hence favourable cases for $C$ are :

(1, 2), (1, 4), (1, 6) ; (4, 1), (4, 3), (4, 5)
(2, 1), (2, 3), (2, 5) ; (5, 2), (5, 4), (5, 6)
(3, 2), (3, 4), (3, 6) ; (6, 1), (6, 3), (6, 5)

*i.e.*, 18 cases in all.

Hence $P(C) = \frac{18}{36} = \frac{1}{2}$.

Cases favourable to the events $A \cap B, A \cap C, B \cap C$ and $A \cap B \cap C$ are given below :

| *Event* | *Favourable cases* |
|---|---|
| $A \cap B$ | (1,1), (1, 3), (1, 5), (3, 1), (3, 3), (3, 5), (5, 1) (5, 3) (5, 5), *i.e.*, 9 in all. |
| $A \cap C$ | (1, 2), (1, 4), (1, 6), (3, 2), (3, 4), (3, 6), (5, 2), (5, 4) (5, 6), *i.e.*, 9 in all. |
| $B \cap C$ | (2, 1), (4, 1), (6, 1) (2, 3), (4, 3), (6, 3), (2, 5), (4, 5), (6, 5), *i.e.*, 9 in all |
| $A \cap B \cap C$ | Nil, because $A \cap B$ implies that sum of points on two dice is even and hence $(A \cap B) \cap C = \phi$ |

$$\therefore \qquad P(A \cap B) = \frac{9}{36} = \frac{1}{4} = P(A) \cdot P(B)$$

$$P(A \cap C) = \frac{9}{36} = \frac{1}{4} = P(A)\, P(C)$$

$$P(B \cap C) = \frac{9}{36} = \frac{1}{4} = P(B)\, P(C)$$

and $\qquad P(A \cap B \cap C) = P(\phi) = 0 \neq P(A)\, P(B)\, P(C)$

Hence the events $A$, $B$ and $C$ are pairwise independent but not mutually independent.

**Example 4·52.** *Let $A_1, A_2, \ldots, A_n$ be independent events and $P\{A_k\} = p_k$. Further, let p be the probability that none of the events occurs; then show that*

$$p \le e^{-\Sigma p_k}$$

**[Agra Univ. M.Sc., 1987]**

**Solution.** We have

$$p = P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n)$$
$$= \prod_{i=1}^{n} P(\bar{A}_i) = \prod_{i=1}^{n} [1 - P(A_i)] = \prod_{i=1}^{n} (1 - p_i)$$

[since $A_i$'s are independent]

$$\leq \prod_{i=1}^{n} e^{-p_i}$$

[$\because 1 - x \leq e^{-x}$ for $0 \leq x \leq 1$ and $0 \leq p_i \leq 1$ ]

$$\Rightarrow \quad p \leq \exp\left[-\sum_{i=1}^{n} p_i\right],$$

as desired.

**Remark.** We have

$$1 - x \leq e^{-x} \text{ for } 0 \leq x \leq 1 \qquad \dots(*)$$

**Proof.** The inequality (*) is obvious for $x = 0$ and $x = 1$. Consider $0 < x < 1$. Then

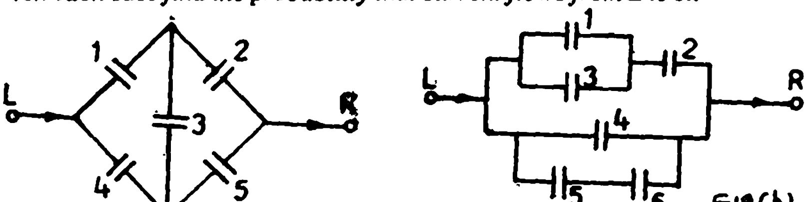
$$\log(1-x)^{-1} = -\log(1-x)$$
$$= \left[x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \dots\right], \qquad \dots(**)$$

the expansion being valid since $0 < x < 1$. Further since $x > 0$, we get from (**)

$$\log(1-x)^{-1} > x$$
$$\Rightarrow \quad -\log(1-x) > x$$
$$\Rightarrow \quad \log(1-x) < -x$$
$$\Rightarrow \quad 1 - x < e^{-x},$$

as desired.

**Example 4·53.** *In the following Fig.(a) and (b) assume that the probability of a relay being closed is P and that a relay is open or closed independently of any other. In each case find the probability that current flows from L to R.*




Fig(a) Fig(b)

**Solution.** Let $A_i$ denote the event that the relay $i$, ($i = 1, 2, \dots, 6$) is closed. Let $E$ be the event that current flows from $L$ to $R$.

*In Fig. (a) the current will flow from L to R if at least one of the circuits from L to R is closed.* Thus for the current to flow from $L$ to $R$ we have the following favourable cases:

(i) $A_1 \cap A_2 = B_1$, (ii) $A_4 \cap A_5 = B_2$,

(iii) $A_1 \cap A_3 \cap A_5 = B_3$, (iv) $A_4 \cap A_3 \cap A_2 = B_4$,

The probability $p_1$ that current flows from $L$ to $R$ is given by

$$p_1 = P(B_1 \cup B_2 \cup B_3 \cup B_4) = \sum_i P(B_i) - \sum_{i<j} P(B_i \cap B_j) + \sum_{i<j<k} P(B_i \cap B_j \cap B_k) - P(B_1 \cap B_2 \cap B_3 \cap B_4) \quad ...(*)$$

Since the relays operate independently of each other, we have

$$P(B_1) = P(A_1 \cap A_2) = P(A_1) . P(A_2) = p . p = p^2$$
$$P(B_2) = P(A_4 \cap A_5) = P(A_4) . P(A_5) = p . p = p^2$$
$$P(B_3) = P(A_1) P(A_3) P(A_5) = p^3$$
$$P(B_4) = P(A_4) P(A_3) P(A_2) = p^3$$

Similarly

$$P(B_1 \cap B_2) = P(A_1 \cap A_2 \cap A_4 \cap A_5) = P(A_1) P(A_2) P(A_4) P(A_5) = p^4$$
$$P(B_1 \cap B_2 \cap B_3) = P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) = p^5$$

and so on. Finally, substituting in (*), we get

$$p_1 = (p^2 + p^2 + p^3 + p^3) - (p^4 + p^4 + p^4 + p^4 + p^4 + p^5) + (p^5 + p^5 + p^5 + p^5) - p^5$$
$$= 2p^2 + 2p^3 - 5p^4 + 2p^5$$

*In Fig. (b).* Arguing as in the above case, the required probability $p_2$ that the current flows from $L$ to $R$ is given by

$$p_2 = P(E_1 \cup E_2 \cup E_3 \cup E_4)$$

where

$$E_1 = A_1 \cap A_2, E_2 = A_3 \cap A_2, E_3 = A_4, E_4 = A_5 \cap A_6$$

$$\therefore \quad p_2 = \sum_i P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) - P(E_1 \cap E_2 \cap E_3 \cap E_4)$$
$$= (p^2 + p^2 + p + p^2) - (p^3 + p^3 + p^4 + p^3 + p^4 + p^3) + (p^4 + p^5 + p^5 + p^5) - p^6$$
$$= p + 3p^2 - 4p^3 - p^4 + 3p^5 - p^6$$

**Matching Problem.** Let us have $n$ letters corresponding to which there exist $n$ envelopes bearing different addresses. Considering various letters being put in various envelopes, *a match is said to occur* if a letter goes into the right envelope. (Alternatively, if in a party there are $n$ persons with $n$ different hats, *a match is said to occur* if in the process of selecting hats *at random*, the $i$th person rightly gets the $i$th hat.)

**A match at the kth position for k=1, 2, ..., n.** Let us first consider the event $A_k$ when a match occurs at the $k$th place. For better understanding let us put the envelopes bearing numbners 1, 2, ..., $n$ in ascending order. When $A_k$ occurs, $k$ th letter goes to the $k$th envelope but $(n-1)$ letters can go to the remaining $(n-1)$ envelopes in $(n-1)!$ ways.

Hence $$P(A_k) = \frac{(n-1)!}{n!} = \frac{1}{n},$$

where $P(A_k)$ denotes the probability of the $k$th match. It is interesting to see that *$P(A_k)$ does not depend on $k$.*

**Example 4·54.** **(a)** *'$n$' different objects $1, 2, \ldots, n$ are distributed at random in $n$ places marked $1, 2, \ldots, n$. Find the probability that none of the objects occupies the place corresponding to its number.* **[Calcutta Univ. B.A.(Stat.Hons.)1986; Delhi Univ. B.Sc.(Maths Hons.), 1990; B.Sc.(Stat.Hons.) 1988]**

*(b) If $n$ letters are randomly placed in correctly addressed envelopes, prove that the probability that exactly $r$ leters are placed in correct envelopes is given by*

$$\frac{1}{r!} \sum_{k=0}^{n-r} (-1)^k \frac{1}{k!}\ ; \ r = 1, 2, \ldots, n$$

**[Bangalore Univ. B.Sc., 1987]**

**Solution (Probability of no match).** Let $E_i$, $(i = 1, 2, \ldots, n)$ denote the event that the $i$th object occupies the place corresponding to its number so that $\overline{E}_i$, is the complementary event. Then the probability '$p$' that none of the objects occupies the place corresponding to its number is given by

$$\begin{aligned} p &= P(\overline{E}_1 \cap \overline{E}_2 \cap \overline{E}_3 \cap \ldots \overline{E}_n) \\ &= 1 - P\ \{\text{at least one of the objects occupies the place corresponding to its number}\} \\ &= 1 - P(E_1 \cup E_2 \cup E_3 \cup \ldots \cup E_n) \\ &= 1 - \Big[\sum_{i=1}^{n} P(E_i) - \sum\sum_{\substack{i,j=1 \\ i<j}}^{n} P(E_i \cap E_j) + \sum\sum\sum_{\substack{i,j,k=1 \\ i<j<k}}^{n} P(E_i \cap E_j \cap E_k) - \ldots \\ &\qquad + (-1)^{n-1} P(E_1 \cap E_2 \cap \ldots \cap E_n)\Big] \qquad \ldots(*) \end{aligned}$$

Now $$P(E_i) = \frac{1}{n}, \ \forall \ i$$

$$\begin{aligned} P(E_i \cap E_j) &= P(E_i)\, P(E_j \mid E_i) \\ &= \frac{1}{n} \cdot \frac{1}{n-1}, \ \forall \ i, j\ (i<j) \\ P(E_i \cap E_j \cap E_k) &= P(E_i)\, P(E_j \mid E_i)\, P(E_k \mid E_i \cap E_j) \\ &= \frac{1}{n} \cdot \frac{1}{n-1} \cdot \frac{1}{n-2}, \ \forall \ i, j, k\ (i<j<k) \end{aligned}$$

and so on. Finally,

$$P(E_1 \cap E_2 \cap E_3 \cap \ldots \cap E_n) = \frac{1}{n} \cdot \frac{1}{n-1} \cdot \frac{1}{n-2} \ldots \frac{1}{2} \cdot 1$$

Substituting in (*), we get

$$p = 1 - \left[ {}^nC_1 \frac{1}{n} - {}^nC_2 \frac{1}{n(n-1)} + {}^nC_3 \frac{1}{n(n-1)(n-2)} - \dots + (-1)^{n-1} \frac{1}{n(n-1)\dots 3 \,.\, 2 \,.\, 1} \right]$$

$$= 1 - \left[ 1 - \frac{1}{2\,!} + \frac{1}{3\,!} - \dots + (-1)^{n-1} \frac{1}{n\,!} \right]$$

$$= \frac{1}{2\,!} - \frac{1}{3\,!} + \frac{1}{4\,!} - \dots + (-1)^n \frac{1}{n\,!}$$

$$= \sum_{k=0}^{n} \frac{(-1)^k}{k\,!}$$

**Remark.** For large $n$,

$$p = 1 - 1 + \frac{1}{2\,!} - \frac{1}{3\,!} + \frac{1}{4\,!} - \dots$$

$$= e^{-1} = 0{\cdot}36787$$

Hence the *probability of at least one match is*

$$1 - p = 1 - \frac{1}{2\,!} + \frac{1}{3\,!} - \dots + \frac{(-1)^n}{n\,!}$$

$$= 1 - \frac{1}{e}, \text{ (for large } n)$$

*(b)* **[Probability of exactly $r$ matches $\{r \le (n-2)\}$]** Let $A_i$, $(i = 1, 2, \dots, n)$ denote the event that $i$th letter goes to the correct envelope. Then the probability that none of the $n$ letters goes to the correct envelope is

$$P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n) = \sum_{k=0}^{n} (-1)^k / k\,! \qquad \dots(**) \ [(c.f. \text{ part } (a)]$$

The probability that each of the '$r$' letters is in the right envelope is $\dfrac{1}{n\,(n-1)\,(n-2)\,\dots\,(n-r+1)}$, and the probability that none of the remaining $(n-r)$ letters goes in the correct envelope is obtained by replacing $n$ by $(n-r)$ in (**) and is thus given by $\displaystyle\sum_{k=0}^{n-r} \frac{(-1)^k}{k\,!}$. Hence by compound probability theorem, the probability that out of $n$ letters exactly $r$ letters go to correct envelopes, (in a specified order), is

$$\frac{1}{n\,(n-1)\,(n-2)\,\dots\,(n-r+1)} \sum_{k=0}^{n-r} \frac{(-1)^k}{k\,!}\,; \quad r \le n-2.$$

Since $r$ letters can go to $n$ envelopes in ${}^nC_r$ mutually exclusive ways, the required probability of exactly $r$ letters going to correct envelopes, (in any order, whatsoever), is given by

$$^nC_r \times \frac{1}{n(n-1)(n-2)\ldots(n-r+1)} \sum_{k=0}^{n-r} \frac{(-1)^k}{k!} = \frac{1}{r!} \sum_{k=0}^{n-r} (-1)^k \frac{1}{k!}$$

**Example 4·55.** *Each of the n urns contains 'a' white balls and 'b' black balls. One ball is transferred from the first urn to the second, then one ball from the latter into the third, and so on. If $p_k$ is the probability of drawing a white ball from the kth urn, show that*

$$p_{k+1} = \frac{a+1}{a+b+1} p_k + \frac{a}{a+b+1}(1-p_k)$$

*Hence for the last urn, prove that*

$$p_n = \frac{a}{a+b}$$ **[Punjab Univ. B.Sc.(Maths Hons.),1988]**

**Solution.** The event of drawing a white ball from the $k$th urn can materialise in the following two ways:

*(i)* The ball transferred from the $(k-1)$th urn is white and then a white ball is drawn from the kth urn.

*(ii)* The ball transferred from the $(k-1)$th urn is black and then a white ball is drawn from the $k$th urn.

The probability of case *(i)* is $p_{k-1} \times \dfrac{a+1}{a+b+1}$,

since the probability of drawing a white ball from the $(k-1)$th urn is $p_{k-1}$ and then the probability of drawing white ball from the $k$th urn is

$$\frac{a+1}{a+b+1}.$$

Since the probability of drawing a black ball from the $(k-1)$th urn is $[1-p_{k-1}]$ and then the probability of drawing a white ball from the $k$th urn is

$$\frac{a}{a+b+1},$$

the probability of case *(ii)* is given by

$$\frac{a}{a+b+1}[1-p_{k-1}]$$

Since the cases *(i)* and *(ii)* are mutually exclusive, we have by addition theorem of probability

$$p_k = \frac{a+1}{a+b+1} p_{k-1} + \frac{a}{a+b+1}[1-p_{k-1}] \qquad \ldots(*)$$

$$\therefore \quad p_k = \frac{1}{a+b+1} p_{k-1} + \frac{a}{a+b+1} \qquad \ldots(1)$$

Replacing $k$ by $k+1$ in (*) we get the required result.

Changing $k$ to $k-1, k-2, \ldots$ and so on, we get

$$p_{k-1} = \frac{1}{a+b+1} p_{k-2} + \frac{a}{a+b+1} \qquad \ldots(2)$$

$$p_{k-2} = \frac{1}{a+b+1}\, p_{k-3} + \frac{a}{a+b+1} \qquad \ldots(3)$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$p_2 = \frac{1}{a+b+1}\, p_1 + \frac{a}{a+b+1} \qquad \ldots(k-1)$$

But $p_1$ = Probability of drawing a white ball from the first urn $= \dfrac{a}{a+b}$.

Multiplying (1) by 1, (2) by $\dfrac{1}{a+b+1}$, (3) by $\left(\dfrac{1}{a+b+1}\right)^2$, ..., and $(k-1)$th equation by $\left(\dfrac{1}{a+b+1}\right)^{k-2}$ and adding, we get

$$p_k = \left(\frac{1}{a+b+1}\right)^{k-1} p_1 + \frac{a}{a+b+1}\left[1 + \frac{1}{a+b+1} + \frac{1}{(a+b+1)^2} + \ldots + \left(\frac{1}{a+b+1}\right)^{k-2}\right]$$

$$= \left(\frac{1}{a+b+1}\right)^{k-1} \times \frac{a}{(a+b)} + \frac{a}{a+b+1}\left[\frac{1-\left(\dfrac{1}{a+b+1}\right)^{k-1}}{\left(1-\dfrac{1}{a+b+1}\right)}\right]$$

$$= \frac{a}{a+b}\left(\frac{1}{a+b+1}\right)^{k-1} + \frac{a}{a+b}\left[1-\left(\frac{1}{a+b+1}\right)^{k-1}\right]$$

$$= \frac{a}{a+b}\left[\left(\frac{1}{a+b+1}\right)^{k-1} + \left\{1-\left(\frac{1}{a+b+1}\right)^{k-1}\right\}\right]$$

$$= \frac{a}{a+b}, \quad (k = 1, 2, \ldots, n)$$

Since the probability of drawing a white ball from the $k$th urn is independent of $k$, we have

$$p_n = \frac{a}{a+b}.$$

**Example 4·56.** *(i) Let the probability $p_n$ that a family has exactly n children be $\alpha p^n$ when $n \geq 1$ and $p_0 = 1 - \alpha p\,(1 + p + p^2 + \ldots)$. Suppose that all sex distributions of n children have the same probability. Show that for $k \geq 1$, the probability that a family contains exactly k boys is $2\alpha \,.\, p^k/(2-p)^{k+1}$.*

*(ii) Given that a family includes at least one boy, show that the probability that there are two or more boys is $p/(2-p)$.*

**Solution.** We are given

$$p_n = P\text{ [that a family has exactly } n \text{ children]}$$
$$= \alpha p^n,\ n \geq 1$$

and $\quad p_0 = 1 - \alpha p\,(1 + p + p^2 + \ldots)$

Let $E_j$ be the event that the number of children in a family is $j$ and let $A$ be the event that a family contains exactly $k$ boys. Then

$$P(E_j) = p_j;\ j = 0, 1, 2, \ldots$$

Now, since each child can have any of the two sex distributions (either boy or girl), the total number of possible distributions for a family to have '$j$' children is $2^j$.

$$\therefore \qquad P(A \mid E_j) = \frac{{}^jC_k}{2^j},\ j \geq k$$

and
$$P(A) = \sum_{j=k}^{\infty} P(E_j)\,P(A \mid E_j) = \sum_{j=k}^{\infty} p_j P(A \mid E_j)$$
$$= \sum_{j=k}^{\infty} \alpha p^j \left[\frac{{}^jC_k}{2^j}\right], \quad j \geq k \geq 1$$
$$= \alpha \sum_{j=k}^{\infty} \left(\frac{p}{2}\right)^j {}^jC_k$$
$$= \alpha \sum_{r=0}^{\infty} {}^{k+r}C_k \left(\frac{p}{2}\right)^{k+r} \qquad \text{[Put } j-k=r\text{]}$$
$$= \alpha \left(\frac{p}{2}\right)^k \sum_{r=0}^{\infty} {}^{k+r}C_r \left(\frac{p}{2}\right)^r \qquad [\because {}^nC_r = {}^nC_{n-r}]$$

We know that

$${}^{-n}C_r = (-1)^r\,.\,{}^{n+r-1}C_r \quad \Rightarrow \quad (-1)^r\,.\,{}^{-n}C_r = {}^{n+r-1}C_r$$
$$\therefore \qquad (-1)^r\,.\,{}^{-(k+1)}C_r = {}^{k+r}C_r$$

Hence

$$P(A) = \alpha \left(\frac{p}{2}\right)^k \sum_{r=0}^{\infty} (-1)^r\,.\,{}^{-(k+1)}C_r \left(\frac{p}{2}\right)^r$$
$$= \alpha \left(\frac{p}{2}\right)^k \sum_{r=0}^{\infty} {}^{-(k+1)}C_r \left(-\frac{p}{2}\right)^r$$
$$= \alpha \left(\frac{p}{2}\right)^k \left(1 - \frac{p}{2}\right)^{-(k+1)}$$
$$= \alpha \left(\frac{p}{2}\right)^k \frac{2^{k+1}}{(2-p)^{k+1}} = \frac{2\alpha p^k}{(2-p)^{k+1}}$$

*(b)* Let $B$ denote the event that a family includes at least one boy and $C$ denote the event that a family has two or more boys. Then

$$P(B) = \sum_{k=1}^{\infty} P\,[\text{family has exactly } k \text{ boys}]$$

$$= \sum_{k=1}^{\infty} \frac{2\,\alpha\, p^k}{(2-p)^{k+1}} = \frac{2\,\alpha}{2-p} \sum_{k=1}^{\infty} \left(\frac{p}{2-p}\right)^k$$

$$= \frac{2\,\alpha}{2-p} \times \frac{p/(2-p)}{1-[p/(2-p)]} = \frac{\alpha\, p}{(1-p)\,(2-p)}$$

$$P(C) = \sum_{k=2}^{\infty} P\,[\text{family has exactly } k \text{ boys}]$$

$$= \sum_{k=2}^{\infty} \frac{2\,\alpha\, p^k}{(2-p)^{k+1}} = \frac{2\,\alpha}{2-p} \sum_{k=2}^{\infty} \left(\frac{p}{2-p}\right)^k$$

$$= \frac{2\,\alpha}{2-p} \cdot \frac{[p/(2-p)]^2}{1-[p/(2-p)]} = \frac{\alpha\, p^2}{(2-p)^2\,(1-p)}$$

Since $C \subset B$ and $B \cap C = C, P(B \cap C) = P(C) \Rightarrow P(B)\,P(C \mid B) = P(C)$

Therefore,

$$P(C \mid B) = \frac{P(C)}{P(B)} = \frac{\alpha\, p^2}{(2-p)^2\,(1-p)} \times \frac{(1-p)\,(2-p)}{\alpha\, p} = \frac{p}{2-p}$$

**Example 4·57.** *A slip of paper is given to person A who marks it either with a plus sign or a minus sign; the probability of his writing a plus sign is 1/3. A passes the slip to B, who may either leave it alone or change the sign before passing it to C. Next C passes the slip to D after perhaps changing the sign. Finally D passes it to a referee after perhaps changing the sign. The referee sees a plus sign on the slip. It is known that B, C and D each change the sign with probability 2/3. Find the probability that A originally wrote a plus.*

**Solution.** Let us define the following events :

$E_1$ : A wrote a plus sign; $E_2$ : A wrote a minus sign

$E$ : The referee observes a plus sign on the slip.

We are given : $P(E_1) = 1/3$, $P(E_2) = 1 - 1/3 = 2/3$

We want $P(E_1 \mid E)$, which by Bayes rule is given by :

$$P(E_1 \mid E) = \frac{P(E_1)\,P(E \mid E_1)}{P(E_1)\,P(E \mid E_1) + P(E_2)\,P(E \mid E_2)} \qquad \text{...(i)}$$

$P(E \mid E_1) = P$ [Referee observes the plus sign given that '$A$' wrote the plus sign on the slip]

$= P$ [(Plus sign was not changed at all) $\cup$ (Plus sign was changed exactly twice in passing from '$A$' to referee through $B$, $C$ and $D$)]

$= P(E_3 \cup E_4)$, (say).

$= P(E_3) + P(E_4)$ ...(ii)

Let $A_1, A_2$ and $A_3$ respectively denote the events that $B$, $C$ and $D$ change the sign on the slip. Then we are given

$P(A_1) = P(A_2) = P(A_3) = 2/3 \quad ; \quad P(\bar{A}_1) = P(\bar{A}_2) = P(\bar{A}_3) = 1/3$

We have

$P(E_3) = P(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3) = P(\bar{A}_1)\, P(\bar{A}_2)\, P(\bar{A}_3) = (1/3)^3 = 1/27$

$$\begin{aligned} P(E_4) &= P[(A_1 A_2 \bar{A}_3) \cup (A_1 \bar{A}_2 A_3) \cup (\bar{A}_1 A_2 A_3)] \\ &= P(A_1 A_2 \bar{A}_3) + P(A_1 \bar{A}_2 A_3) + P(\bar{A}_1 A_2 A_3) \\ &= P(A_1)\, P(A_2)\, P(\bar{A}_3) + P(A_1)\, P(\bar{A}_2)\, P(A_3) + P(\bar{A}_1)\, P(A_2)\, P(A_3) \\ &= \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9} \end{aligned}$$

Substituting in *(ii)* we get

$$P(E \mid E_1) = \frac{1}{27} + \frac{4}{9} = \frac{13}{27} \qquad \ldots (iii)$$

Similarly,

$P(E \mid E_2) = P$ [Referee observes the plus sign given that '$A$' wrote minus sign on the slip]

$= P$ [(Minus sign was changed exactly once) $\cup$ (Minus sign was changed thrice)]

$= P(E_5 \cup E_6)$, (say),

$= P(E_5) + P(E_6) \qquad \ldots (iv)$

$$\begin{aligned} P(E_5) &= P[(A_1 \bar{A}_2 \bar{A}_3) \cup (\bar{A}_1 A_2 \bar{A}_3) \cup (\bar{A}_1 \bar{A}_2 A_3)] \\ &= P(A_1)\, P(\bar{A}_2)\, P(\bar{A}_3) + P(\bar{A}_1)\, P(A_2)\, P(\bar{A}_3) + P(\bar{A}_1)\, P(\bar{A}_2)\, P(A_3) \\ &= \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9} \end{aligned}$$

$$P(E_6) = P(A_1 A_2 A_3) = P(A_1)\, P(A_2)\, P(A_3) = \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{8}{27}$$

Substituting in *(iv)* we get :

$$P(E \mid E_2) = \frac{2}{9} + \frac{8}{27} = \frac{14}{27}. \qquad \ldots (v)$$

Substituting from *(iii)* and *(v)* in *(i)* we get :

$$P(E_1 \mid E) = \frac{\frac{1}{3} \times \frac{13}{27}}{\frac{1}{3} \times \frac{13}{27} + \frac{2}{3} \times \frac{14}{27}} = \frac{13}{13 + 28} = \frac{13}{41}$$

**Example 4·58.** *Three urns of the same appearance have the following proportion of balls.*

| | | | |
|---|---|---|---|
| *First urn* | : | *2 black* | *1 white* |
| *Second Urn* | : | *1 black* | *2 white* |
| *Third urn* | : | *2 black* | *2 white* |

*One of the urns is selected and one ball is drawn. It turns out to be white. What is the probability of drawing a white ball again, the first one not having been returned?*

**Solution.** Let us define the events:

$E_i$ = The event of selection of $i$th urn, $(i = 1,2,3)$

and $A$ = The event of drawing a white ball.

Then

$$P(E_1) = P(E_2) = P(E_3) = 1/3$$

and $P(A \mid E_1) = 1/3, P(A \mid E_2) = 2/3$ and $P(A \mid E_3) = 1/2$

Let $C$ denote the future event of drawing another white ball from the urns. Then

$$P(C \mid E_1 \cap A) = 0, P(C \mid E_2 \cap A) = 1/2, \text{ and } P(C \mid E_3 \cap A) = 1/3$$

$$\therefore \quad P(C \mid A) = \frac{\sum_{i=1}^{3} P(E_i)\, P(A \mid E_i)\, P(C \mid E_i \cap A)}{\sum_{i=1}^{3} P(E_i)\, P(A \mid E_i)}$$

$$= \frac{\frac{1}{3}\cdot\frac{1}{3}\cdot 0 + \frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{2} + \frac{1}{3}\cdot\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{3}\cdot\frac{1}{3} + \frac{1}{3}\cdot\frac{2}{3} + \frac{1}{3}\cdot\frac{1}{2}} = \frac{1}{3}$$

## MISCELLANEOUS EXERCISE ON CHAPTER IV

1. Probabilities of occurrence of $n$ independent events $E_1, E_2, ..., E_n$ are $p_1, p_2, ..., p_n$ respectively. Find the probability of occurrence of the compound event in which $E_1, E_2, ..., E_r$ occur and $E_{r+1}, E_{r+2}, ..., E_n$ do not occur.

**Ans.** $\prod_{i=1}^{r} p_i \times \prod_{i=r+1}^{n} (1 - p_i)$

2. Prove that for any integer $m \geq 1$,

(a) $P\left(\bigcap_{i=1}^{m} A_i\right) \leq P(A_i) \leq P\left(\bigcup_{i=1}^{m} A_i\right) \leq \sum_{i=1}^{m} P(A_i)$

(b) $P\left(\bigcap_{i=1}^{m} A_i\right) \geq 1 - \sum_{i=1}^{m} P(\overline{A}_i)$

3. Establish the inequalities :

$$P(A \cap B \cap C) \leq P(A \cap B) \leq P(A \cup B) \leq P(A \cup B \cup C) \leq P(A) + P(B) + P(C)$$

4. Let $A_1, A_2, ..., A_n$ be mutually independent events with $P(A_k) = p_k$, $k = 1, 2, ..., n$.

Let $p$ be the probability that none of the events $A_1, A_2, ..., A_n$ occurs. Show that

$$p = \prod_{k=1}^{n} (1 - p_k) \leq \exp\left\{-\sum_{k=1}^{n} p_k\right\}$$

Use the above relation to compute the probability that in six tosses of a fair die, no "aces are obtained". Compare this with the upper bound given above. Show that if each $p_k$ is small compared with $n$, the upper bound is a good approximation.

5. $A$ and $B$ play a match, the winner being the one who first wins two games in succession, no games being drawn. Their respective chances of winning a particular game are $p : q$. Find

(i) $A$'s initial chance of winning.

(ii) $A$'s chance of winning after having won the first game.

6. A carpenter has a tool chest with two compartments, each one having a lock. He has two keys for each lock, and he keeps all four keys in the same ring. His habitual procedure in opening a compartment is to select a key at random and try it. If it fails, he selects one of the remaining three and tries it and so on. Show that the probability that he succeeds on the first, second and third try is 1/2, 1/3, 1/6 respectively. **(Lucknow Univ. B.Sc., 1990)**

7. Three players $A$, $B$ and $C$ agree to play a series of games observing the following rules : two players participate in each game, while third is idle, and the game is to be won by one of them. The loser in each game quits and his place in the next game is taken by the player who was idle. The player who succeeds in winning over both of his opponents without interruption wins the whole series of games.

Supposing the probability for each player to win a single game is 1/2, and that the first game is played by $A$ and $B$, find the probability for $A$, $B$ and $C$ respectively to win the whole series if the number of games is unlimited.

**Ans.** 5/14, 5/14, 2/7

8. In a certain group of mathematicians, 60 per cent have insufficient background of modern Algebra, 50 per cent have inadequate knowledge of Mathematical Statistics and 80 per cent are in either one or both of the two categories. What is the percentage of people who know Mathematical Statistics among those who have a sufficient background of Modern Algebra? **(Ans. 0·50)**

9. (a) If $A$ has $(n+1)$ and $B$ has $n$ fair coins, which they flip, show that the probability that $A$ gets more heads than $B$ is $\frac{1}{2}$.

(b) A student is given a column of 10 dates and column of 10 events and is asked to match the correct date to each event. He is not allowed to use any item more than once. Consider the case where the student knows how to match four of the items but he is very doubtful of the remaining six. He decides to match these at random. Find the probabilities that he will correctly match (i) all the items, (ii) at least seven of the items, and (iii) at least five.

**Ans.** (a) $\dfrac{1}{6\,!}$, (b) $\dfrac{10}{6\,!}$, (c) $1-\dfrac{1}{6\,!}$

10. An astrologer claims that he can predict before birth the sex of a baby just to be born. Suppose that the astrologer has no real power but he tosses a coin just

once before every birth and if the head turns up he predicts a boy for that birth and if the tail turns up he predicts a girl. Let $p$ be the probability of the event that at a certain birth a male child is born, and $p'$ the probability of a head turning up in a single toss with astrologer's coin. Find the probability of a correct prediction and that of at least one correct prediction in $n$ predictions.

**11.** From a pack of 52 cards an even number of cards is drawn. Show that the probability of half of these cards being red is

$$[52\,!/(26\,!)^2 - 1]\,/\,(2^{51} - 1)$$

**12.** A sportsman's chance of shooting an animal at a distance $r\ (> a)$ is $a^2/r^2$. He fires when $r = 2a$, and if he misses he reloads and fires when $r = 3a, 4a\ldots$ If he misses at distance $na$, the animal escapes. Find the odds against the sportsman.

**Ans.** $n + 1 : n - 1$

**Hint.** $P$ [Sportsman shoots at a distance $ia$] $= \dfrac{a^2}{(ia)^2} = \dfrac{1}{i^2}$

$\Rightarrow$ $P$ [Sportsman misses the shot at a distance $ia$] $= 1 - \dfrac{1}{i^2}$

$$\therefore \quad P\,[\text{Animal escapes}] = \prod_{i=2}^{n}\left(1 - \frac{1}{i^2}\right) = \prod_{i=2}^{n}\left[\left(\frac{i-1}{i}\right)\left(\frac{i+1}{i}\right)\right]$$

$$= \prod_{i=2}^{n}\left(\frac{i-1}{i}\right)\prod_{i=2}^{n}\left(\frac{i+1}{i}\right) = \frac{n+1}{2n}$$

$$\text{Required ratio} = \frac{n+1}{2n} : \left(1 - \frac{n+1}{2n}\right) = (n+1) : (n-1)$$

**13.** *(a)* Pataudi, the captain of the Indian team, is reported to have observed the rule of calling 'heads' every time the toss was made during the five matches of the Test series with the Australian team. What is the probability of his winning the toss in all the five matches?

**Ans.** $(1/2)^5$

How will the probability be affected if

*(i)* he had made a rule of tossing a coin privately to decide whether to call "heads" or "tails" on each occasion.

*(ii)* the factors determining his choice were not predetermined but he called out whatever occurred to him on the spur of the moment?

*(b)* A lot contains 50 defective and 50 non-defective bulbs. Two bulbs are drawn at random one at a time, with replacement. The events $A, B, C$ are defined as

$A$ = {The first bulb is defective}
$B$ = {The second bulb is non-defective}
$C$ = {The two bulbs are both defective or both non-defective}

Determine whether

*(i)* $A, B, C$ are pairwise independent,

*(ii)* $A, B, C$ are independent.

**14.** $A, B$ and $C$ are three urns which contain 2 white, 1 black, 3 white, 2 black and 2 white and 2 black balls, respectively. One ball is drawn from urn $A$ and put into the urn $B$; then a ball is drawn from urn $B$ and put into the urn $C$. Then a ball is drawn from urn $C$. Find the probability that the ball drawn is white.

**Ans.** 4/15.

**15.** An urn contains $a$ white and $b$ black balls and a series of drawings of one ball at a time is made, the ball removed being retrurned to the urn immediately after the next drawing is made. If $p_n$ denotes the probability that the $n$th ball drawn is black, show that

$$p_n = (b - p_{n-1}) / (a + b - 1).$$

Hence find $p_n$.

**16.** A person is to be tested to see whether he can differentiate between the taste of two brands of cigarettes. If he cannot differentiate, it is assumed that the probability is one-half that he will identify a cigarette correctly. Under which of the following two procedures is there less chance that he will make all correct identifications when he actually cannot differentiate between the two brands?

*(i)* The subject is given four pairs each containing both brands of cigarettes (this is known to the subject), he must identify for each pair which cigarette represents each brand.

*(ii)* The subject is given eight cigarettes and is told that the first four are of one brand and the last four of the other brand.

How do you explain the difference in results despite the fact that eight cigarettes are tested in each case?

**Ans.** *(i)* 1/16 *(ii)* 1/2

**17.** (Sampling with replacement). A sample of size $r$ is taken from a population of $n$ people. Find the probability $U_r$ that $N$ given people will be included in the sample.

**Ans.** $U_r = \sum_{m=0}^{N} (-1)^m \binom{N}{m} \left(1 - \frac{m}{n}\right)^r$

**18.** In a lottery $m$ tickest are drawn at a time out of the total number of $n$ tickets, and returned before the next drawing is made. Show that the chance that in $k$ drawings, each of the numbers 1, 2, 3, ..., $n$ will appear at least once is given by

$$P_k = 1 - \binom{n}{1}\left(1 - \frac{m}{n}\right)^k + \binom{n}{2}\left(1 - \frac{m}{n}\right)^k \left(1 - \frac{m}{n-1}\right)^k - \ldots$$

**[Nagpur Univ. M.Sc. 1987]**

**19.** In a certain book of $N$ pages, no page contains more than four errors, $n_1$ of them contain one error, $n_2$ contain two errors, $n_3$ contain three errors and $n_4$ contain four errors. Two copies of the book are opened at any two given pages. Show the probability that the number of errors in these two pages shall not exceed five is

$$1 - \frac{1}{N^2}\,(n_3^2 + n_4^2 + 2n_2 n_4 + 2n_3 n_4)$$

**Hint.** Let $E_i$ I : the event that a page of first book contains $i$ errors.

and $E_i$ II : the event that a page of second book contains $i$ errors.

$P$ (No. of errors in the two pages shall not exceed 5)

$$= 1 - P\,[E_2\,\mathrm{I}\,E_4\,\mathrm{II} + E_3\,\mathrm{I}\,E_4\,\mathrm{II} + E_4\,\mathrm{I}\,E_4\,\mathrm{II} + E_3\,\mathrm{I}\,E_3\,\mathrm{II} + E_4\,\mathrm{I}\,E_3\,\mathrm{II} + E_4\,\mathrm{I}\,E_2\,\mathrm{II}\,].$$

**20.** *(a)* Of three independent events, the chance that the first only should happens is $a$, the chance of the second only is $b$ and the chance of the third only is $c$. Show that the independent chances of the three events are respectively :

$$\frac{a}{a+x},\ \frac{b}{b+x},\ \frac{c}{c+x}$$

where $x$ is the root of the equation

$$(a+x)(b+x)(c+x) = x^2$$

**Hint.** $P(E_1 \cap \bar{E}_2 \cap \bar{E}_3) = P(E_1)\,[1 - P(E_2)]\,[1 - P(E_3)] = a$ ...(*)

$P(\bar{E}_1 \cap E_2 \cap \bar{E}_3) = [1 - P(E_1)]\,P(E_2)\,[1 - P(E_3)] = b$ ...(**)

$P(\bar{E}_1 \cap \bar{E}_2 \cap E_3) = [1 - P(E_1)]\,[1 - P(E_2)]\,P(E_3) = c$ ...(***)

Multiplying (*), (**) and (***), we get

$$P(E_1)\,P(E_2)\,P(E_3)\,x^2 = abc,$$

where $x = [1 - P(E_1)]\,[1 - P(E_2)]\,[1 - P(E_3)]$

Multiplying (*) by $[1 - P(E_1)]$, we get

$$P(E_1) = \frac{a}{a+x},\ \text{and so on.}$$

*(b)* Of three independent events, the probability that the first only should happens is 1/4, the probability that the second only should happen is 1/8, and the probability that the third only should happen is 1/12. Obtain the unconditional probabilities of the three events.

**Ans.** 1/2, 1/3, 1/4.

*(c)* A total of $n$ shells are fired at a target. The probability of the $i$th shell hitting the target is $p_i$ ; $i = 1, 2, 3, ..., n$. Assuming that the $n$ firings are $n$ mutually independent events, find the probability that at least two shells out of $n$ hit the target. **[Calcutta Univ. B.Sc.(Maths Hons.), 1988]**

*(d)* An urn contains $M$ balls numbered 1 to $M$, where the first $K$ balls are defective and the remaining $M - K$ are non-defective. A sample of $n$ balls is drawn from the urn. Let $A_k$ be the event that the sample of $n$ balls contains exactly $k$ defectives. Find $P(A_k)$ when the sample is drawn *(i)* with replacement and, *(ii)* without replacement. **[Delhi Univ. B.Sc. (Maths Hons.), 1989]**

21. For three independent events $A$, $B$ and $C$, the probability for $A$ to occur is $a$, the probability that $A$, $B$ and $C$ will not occur is $b$, and the probability that at least one of the three events will not occur is $c$. If $p$ denotes the probability that $C$ occurs but neither $A$ nor $B$ occurs, prove that $p$ satisfies the quadratic equation

$$ap^2 + [ab - (1-a)(a+c-1)]\,p + b\,(1-a)(1-c) = 0$$

and hence deduce that $c > \dfrac{(1-a)^2 + ab}{(1-a)}$

Further show that the probability of occurrence of $C$ is $p/(p+b)$, and that of $B$'s happening is $(1-c)(p+b)/ap$.

**Hint.** Let $P(A) = x$, $P(B) = y$ and $P(C) = z$

Then $\quad x = a,\ (1-x)(1-y)(1-z) = b,\ 1 - xyz = c$

and $\quad p = z(1-x)(1-y)$

Elimination of $x$, $y$ and $z$ gives quadratic equation in $p$.

22. *(a)* The chance of success in each trial is $p$. If $p_k$ is the probability that there are even number of successes in $k$ trials, prove that

$$p_k = p + p_{k-1}(1-2p)$$

Deduce that $\quad p_k = \frac{1}{2}[1 + (1-2p)^k]$

*(b)* If a day is dry, the conditional probability that the following day will also be dry is $p$; if a day is wet, the conditional probability that the following day will be dry is $p'$. If $u_n$ is the probability that the $n$th day will be dry, prove that

$$u_n - (p - p')\,u_{n-1} - p' = 0\ ;\ n \geq 2$$

If the first day is dry, $p = 3/4$ and $p' = 1/4$, find $u_n$.

23. There are $n$ similar biased dice such that the probability of obtaining a 6 with each one of them is the same and equal to $p$. If all the dice are rolled once, show that $p_n$, the probability that an odd number of 6's is obtained satisfies the difference equation

$$p_n + (2p-1)\,p_{n-1} = p$$

and hence derive an explicit expression for $p_n$.

**Ans.** $p_n = \frac{1}{2}[1 + (1-2p)^n]$

24. Suppose that each day the weather can be uniquely classified as 'fine' or 'bad'. Suppose further that the probability of having fine weather on the last day of a certain year is $P_0$ and we have the probability $p$ that the weather on an arbitrary day will be of the same kind as on the preceding day. Let the probability of having fine weather on the $n$th day of the following year be $P_n$. Show that

$$P_n = (2p-1)\,P_{n-1} + (1-p)$$

Deduce that

$$P_3 = (2p-1)^3\left(P_0 - \frac{1}{2}\right) + \frac{1}{2}$$

25. A closet contains $n$ pairs of shoes. If $2r$ shoes are chosen at random (with $2r < n$), what is the probability that there will be *(i)* no complete pair,

*(ii)* exactly one complete pair, *(iii)* exactly two complete pairs among them?

**Hint.** *(i)* $P(\text{no complete pair}) = \binom{n}{2r} 2^{2r} \div \binom{2n}{2r}$

*(ii)* $P(\text{exactly one complete pair}) = n \binom{n-1}{2r-2} 2^{2r-2} \div \binom{2n}{2r}$

and *(iii)* $P(\text{exactly two complete pairs}) = \binom{n}{2}\binom{n-2}{2r-4} 2^{2r-4} \div \binom{2n}{2r}$

26. Show that the probability of getting no right pair out of $n$, when the left foot shoes are paired randomly with the rigth foot shoes, is the sum of the first $(n+1)$ terms in the expansion of $e^{-1}$.

27. *(a)* In a town consisting of $(n+1)$ inhabitants, a person narrates a rumour to a second person, who in turn narrates it to a third person, and so on. At each step the recipient of the rumour is chosen at random from the $n$ available persons, excluding the narrator himself. Find the probability that the rumour will be told $r$ times without:

*(i)* returning to the originator,

*(ii)* being narrated to any person more than once.

*(b)* Do the above problem when, at each step the rumour is told by one person to a gathering of $N$ randomly chosen people.

**Ans.** *(a)* *(i)* $\dfrac{n(n-1)^{r-1}}{n^r} = \left(1 - \dfrac{1}{n}\right)^{r-1}$ ; *(ii)* $\dfrac{n(n-1)(n-2)\ldots(n-r+1)}{n^r}$

*(b)* *(i)* $\left(1 - \dfrac{N}{n}\right)^{r-1}$ ; *(ii)* $\dfrac{\binom{n}{rN}}{\left[\binom{n}{N}\right]^r}$

28. What is the probability that *(i)* the birthdays of twelve people will fall in twelve different calendar months (assume equal probabilities for the twelve months) and *(ii)* the birthdays of six people will fall in exactly two calendar months?

**Hint.** *(i)* The birthday of the first person, for instance, can fall in 12 different ways and so for the second, and so on.

$\therefore$ The total number of cases $= 12^{12}$.

Now there are 12 months in which the birthday of one person can fall and 11 months in which the birthday of the second person can fall and 10 months for another third person, and so on.

$\therefore$ The total number of favourable cases $= 12.11.10\ldots3.2.1$

Hence the required probability $= \dfrac{12!}{12^{12}}$

*(ii)* The total number of ways in which the birthdays of 6 persons can fall in any of the month $= 12^6$.

$\therefore$ The required probability $= \dfrac{\binom{12}{2}(2^6 - 2)}{12^6}$

29. An elevator starts with 7 passengers and stops at 10 floors. What is the probability $p$ that no two passengers leave at the same floor?

[Delhi Univ. M.C.A., 1988]

30. A bridge player knows that his two opponents have exactly five hearts between two of them. Each opponent has thirteen cards. What is the probability that there is three-two split on the hearts (that is one player has three hearts and the other two)? [Delhi Univ. B.Sc.(Maths Hons.), 1988]

31. An urn contains 2 white and 2 black balls. A ball is drawn at random. If it is white, it is not replaced into the urn. Otherwise it is replaced along with another ball of the same colour. The process is repeated. Find the probability that the third ball drawn is black. [Burdwan Univ. B.Sc. (Hons.), 1990]

Ans. $\dfrac{23}{30}$

32. There is a series of $n$ urns. In the $i$th urn there are $i$ white and $(n-i)$ black balls, $i=1, 2, 3, ..., k$. One urn is chosen at random and 2 balls are drawn from it. Both turn out to be white. What is the probability that the $j$th urn was chosen, where $j$ is a particular number between 3 and $n$.

**Hint.** Let $E_j$ denote the event of selection of $j$th urn, $j=3, 4, ..., n$ and $A$ denote the event of drawing of 2 white balls, then

$$P(A \mid E_j)=\left(\frac{j}{n}\right)\left(\frac{j-1}{n-1}\right), \; P(E_j)=\frac{1}{n}, \; P(A)=\sum_{i=1}^{n}\frac{1}{n}\left(\frac{i}{n}\right)\left(\frac{i-1}{n-1}\right)$$

$$\therefore \qquad P(E_j \mid A)=\frac{\frac{1}{n}\left(\frac{j}{n}\right)\left(\frac{j-1}{n-1}\right)}{\sum\limits_{i=1}^{n}\left(\frac{1}{n}\right)\left(\frac{i}{n}\right)\left(\frac{i-1}{n-1}\right)}$$

33. There are $(N+1)$ identical urns marked $0, 1, 2, ..., N$ each of which contains $N$ white and red balls. The $k$th urn contains $k$ red and $N-k$ white balls, $(k=0, 1, 2, ... N)$. An urn is chosen at random and $n$ random drawings of a ball are made from it, the ball drawn being replaced after each draw. If the balls drawn are all red, show that the probability that the next drawing will also yield a red ball is approximately $(n+1)(n+2)$ when $N$ is large.

34. A printing machine can print $n$ letters, say $\alpha_1, \alpha_2, ..., \alpha_n$. It is operated by electrical impulses, each letter being produced by a different impulse. Assume that $p$ is the constant probability of printing the correct letter and the impulses are independent. One of the $n$ impulses, chosen at random, was fed into the machine twice and both times the letter $\alpha_1$ was printed. Compute the probability that the impulse chosen was meant to print $\alpha_1$. [Delhi Univ. M.Sc.(Stat.), 1981]

Ans. $(n-1)p^2/(np^2-2p+1)$

35. Two players $A$ and $B$ agree to contest a match consisting of a series of games, the match to be won by the player who first wins three games, with the provision that if the players win two games each, the match is to continue until it

is won by one player winning two games more than his opponent. The probability of $A$ winning any given game is $p$, and the games cannot be drawn.

*(i)* Prove that $f(p)$, the initial probability of $A$ winning the match is given by:

$$f(p) = p^3 (4 - 5p + 2p^2)/(1 - 2p + 2p^2)$$

*(ii)* Show that the equation $f(p) = p$ has five real roots, of which three are admissible values of $p$. Find these three roots and explain their significance.

[Civil Services (Main), 1986]

36. Two players $A$ and $B$ start playing a series of games with Rs. $a$ and $b$ respectively. The stake is Re. 1 on a game and no game can be drawn. If the probability of $A$ winning any game is a constant $p$, find the initial probability of his exhausting the funds of $B$ or his own. Also show that if the resources of $B$ are unlimited then

*(i)* $A$ is certain to be ruined if $p = \frac{1}{2}$, and

*(ii)* $A$ has an even chance of escaping ruin if $p = 2^{1/a}/(1 + 2^{1/a})$.

**Hint.** Let $u_n$ be the probability of $A$'s final win when he has Rs. $n$.

Then $\quad u_n = pu_{n+1} + (1-p)u_{n-1}$ where $u_o = 0$ and $u_{a+b} = 1$

$$\therefore \quad u_{n+1} - u_n = \left(\frac{1-p}{p}\right)\left(u_n - u_{n-1}\right)$$

Hence $\quad u_{n+1} - u_n = \left(\frac{1-p}{p}\right)^n u_1$, by repeated application,

so that $$u_n = u_1\left[1 - \left(\frac{1-p}{p}\right)^n\right] \Big/ \left[1 - \left(\frac{1-p}{p}\right)\right]$$

Hence using $u_{a+b} = 1$, $$u_n = \left[1 - \left(\frac{1-p}{p}\right)^n\right] \Big/ \left[1 - \left(\frac{1-p}{p}\right)^{a+b}\right]$$

$\therefore$ Initial probability of $A$'s win is $u_a = \dfrac{p^a - (1-p)^a}{p^{a+b} - (1-p)^{a+b}} \cdot p^b$

Probability of $A$'s ruin $= 1 - u_a$.

For $p = \frac{1}{2}$, $u_a = \dfrac{a}{a+b} \to 0$ as $b \to \infty$ and for $p \neq \frac{1}{2}$, $u_a = \frac{1}{2}$ if $p = 2^{1/a}/(1 + 2^{1/a})$.

37. In a game of skill a player has probability 1/3, 5/12 and 1/4 of scoring 0, 1 and 2 points respectively at each trial, the game terminating on the first realization of a zero score at a trial. Assuming that the trials are independent, prove that the probability of the player obtaining a total score of $n$ points is

$$u_n = \frac{3}{13}\left(\frac{3}{4}\right)^n + \frac{4}{39}\left(-\frac{1}{3}\right)^n$$

**Hint.** Event can materialize in the two mutually exclusive ways:

*(i)* at the $(n-1)$th trial, a score of $(n-1)$ points is obtained and a score of 1 point is obtained at the $n$th trial.

*(ii)* at the $(n-2)$th trial, a score of $(n-2)$ points is obtained and a score of 2 points is obtained at the last two trials.

Hence $u_n = \frac{5}{12} u_{n-1} + \frac{1}{4} u_{n-2}$ where $u_0 = \frac{1}{3}$, $u_1 = \frac{1}{3} \cdot \frac{5}{12} = \frac{5}{36}$

Also $u_n = \left( \frac{3}{4} - \frac{1}{3} \right) u_{n-1} + \frac{1}{4} u_{n-2} \Rightarrow u_n + \frac{1}{3} u_{n-1} = \frac{3}{4} \left( u_{n-1} + \frac{1}{3} u_{n-2} \right)$

This equation can be solved as a homogeneous difference equation of second order with the initial conditions

$$u_0 = \frac{1}{3}, \; u_1 = \frac{1}{3} \cdot \frac{5}{12} = \frac{5}{36}$$

**38.** The following weather forecasting is used by an amateur forecaster. Each day is classified as 'dry' or 'wet' and the probability that any given day is same as the preceding one is assumed to be a constant $p$, $(0 < p < 1)$. Based on past records, it is supposed that January 1 has a probability $\beta$ of being dry. Letting

$\beta_n =$ Probability that $n$th day of the year is dry, obtain an expression for $\beta_n$ in terms of $\beta$ and $p$. Also evaluate $\lim\limits_{n \to \infty} \beta_n$.

**Hint.** $\beta_n = p \cdot \beta_{n-1} + (1-p)(1-\beta_{n-1})$

$\Rightarrow$ $\beta_n = (2p-1)\beta_{n-1} + (1-p)$ ; $n = 2, 3, 4, \ldots$

**Ans.** $\beta_n = (2p-1)^{n-1}(\beta - ½) + ½$ ; $\lim\limits_{n \to \infty} \beta_n = ½$

**39.** Two urns contain respectively '$a$ white and $b$ black' and '$b$ white and $a$ black' balls. A series of drawings is made according to the following rules:

*(i)* Each time only one ball is drawn and immediately returned to the same urn it came from.

*(ii)* If the ball drawn is white, the next drawing is made from the first urn.

*(iii)* If it is black, the next drawing is made from the second urn.

*(iv)* The first ball drawn comes from the first urn.

What is the probability that $n$th ball drawn will be white?

**Hint.** $p_r = P$ [Drawing a white ball at the $r$th draw].

$$p_r = \frac{a}{a+b} p_{r-1} + \frac{b}{a+b} \cdot \left( 1 - p_{r-1} \right)$$

$$\Rightarrow \quad p_r = \frac{a-b}{a+b} \cdot p_{r-1} + \frac{b}{a+b}$$

**Ans.** $p_n = \frac{1}{2} + \frac{1}{2} \left( \frac{a-b}{a+b} \right)^n$

**40.** If a coin is tossed repeatedly, show that the probability of getting $m$ heads before $n$ tails is :

$$\frac{1}{2^{m+n-1}} \sum_{i=m}^{m+n-1} {}^{m+n-1}C_i.$$

[Burdwan Univ. (Maths Hons.), 1991]

## OBJECTIVE TYPE QUESTIONS

**I.** Find out the correct answer from group $Y$ for each item of group $X$?

| Group X | Group Y |
|---|---|
| (a) At least one of the events $A$ or $B$ occurs. | (i) $(\bar{A} \cap B) \cup (A \cap \bar{B}) \cup (\bar{A} \cap \bar{B})$ |
| (b) Neither $A$ nor $B$ occurs. | (ii) $(A \cup B) - (A \cap B)$ |
| (c) Exactly one of the events $A$ or $B$ occurs. | (iii) $A \subset B$ |
| | (iv) $B \not\subset A$ |
| (d) If event $A$ occurs, so does $B$. | (v) $[A - (A \cap B)] \cup [B - (A \cap B)]$ |
| (e) Not more than one of the events $A$ or $B$ occur. | (vi) $A \cap \bar{B}$ |
| | (vii) $1 - (A \cup \bar{B})$ |
| | (viii) $A \cup B$ |
| | (ix) $1 - (A \cup B)$ |

**II.** Match the correct expression of probabilities on the left:

| | |
|---|---|
| (a) $P(\phi)$, where $\phi$ is null set | (i) $1 - P(A)$ |
| (b) $P(A \mid B) P(B)$ | (ii) $P(A \cap B)$ |
| (c) $P(\bar{A})$ | (iii) $P(A) - P(A \cap B)$ |
| (d) $P(\bar{A} \cap \bar{B})$ | (iv) $0$ |
| (e) $P(A \sim B)$ | (v) $1 - P(A) - P(B) + P(A \cap B)$ |
| | (vi) $P(A) + P(B) - P(A \cap B)$. |

**III.** Given that $A$, $B$ and $C$ are mutually exclusive events, explain why the following are not permissible assignments of probabilities:

(i) $P(A) = 0{\cdot}24$, $P(B) = 0{\cdot}4$ and $P(A \cup C) = 0{\cdot}2$

(ii) $P(A) = 0{\cdot}4$, $P(B) = 0{\cdot}61$

(iii) $P(A) = 0{\cdot}6$, $P(A \cap \bar{B}) = 0{\cdot}5$

**IV.** In each of the following, indicate whether events $A$ and $B$ are: *(i)* independent, *(ii)* mutually exclusive, *(iii)* dependent but not mutually exclusive.

(a) $P(A \cap B) = 0$ (b) $P(A \cap B) = 0{\cdot}3$, $P(A) = 0{\cdot}45$

(c) $P(A \cup B) = 0{\cdot}85$, $P(A) = 0{\cdot}3$, $P(B) = 0{\cdot}6$

(d) $P(A \cup B) = 0{\cdot}70$, $P(A) = 0{\cdot}5$, $P(B) = 0{\cdot}4$

(e) $P(A \cup B) = 0{\cdot}90$, $P(A \mid B) = 0{\cdot}8$, $P(B) = 0{\cdot}5$.

**V.** Give the correct label as answer like $a$ or $b$ etc., for the following questions:

*(i)* The probability of drawing any one spade card from a pack of cards is

(a) $\frac{1}{52}$ (b) $\frac{1}{13}$ (c) $\frac{4}{13}$ (d) $\frac{1}{4}$

*(ii)* The probability of drawing one white ball from a bag containing 6 red, 8 black, 10 yellow and 1 green balls is

(a) $\frac{1}{25}$ (b) $0$ (c) $1$ (d) $\frac{24}{25}$ (e) $\frac{15}{20}$

*(iii)* A coin is tossed three times in succession, the number of sample points in sample space is

*(a)* 6 *(b)* 8 *(c)* 3

*(iv)* In the simultaneous tossing of two perfect coins, the probability of having at least one head is

*(a)* $\frac{1}{2}$ *(b)* $\frac{1}{4}$ *(c)* $\frac{3}{4}$ *(d)* 1

*(v)* In the simultaneous tossing of two perfect dice, the probability of obtaining 4 as the sum of the resultant faces is

*(a)* $\frac{4}{12}$ *(b)* $\frac{1}{12}$ *(c)* $\frac{3}{12}$ *(d)* $\frac{2}{12}$

*(vi)* A single letter is selected at random from the word 'probability'. The probability that it is a vowel is

*(a)* $\frac{3}{11}$ *(b)* $\frac{2}{11}$ *(c)* $\frac{4}{11}$ *(d)* 0

*(vii)* An urn contains 9 balls, two of which are red, three blue and four black. Three balls are drawn at random. The chance that they are of the same colour is

*(a)* $\frac{5}{84}$ *(b)* $\frac{3}{9}$ *(c)* $\frac{3}{7}$ *(d)* $\frac{7}{17}$

*(viii)* A number is chosen at random among the first 120 natural numbers. The probability of the number chosen being a multiple of 5 or 15 is

*(a)* $\frac{1}{5}$ *(b)* $\frac{1}{8}$ *(c)* $\frac{1}{16}$

*(ix)* If $A$ and $B$ are mutually exclusive events, then

*(a)* $P(A \cup B) = P(A) \,.\, P(B)$

*(b)* $P(A \cup B) = P(A) + P(B)$, *(c)* $P(A \cup B) = 0$.

*(x)* If $A$ and $B$ are two independent events, the probability that both $A$ and $B$ occur is $\frac{1}{8}$ and the probability that neither of them occurs is $\frac{3}{8}$. The probability of the occurrence of $A$ is :

*(a)* $\frac{1}{2}$, *(b)* $\frac{1}{3}$, *(c)* $\frac{1}{4}$, *(d)* $\frac{1}{5}$.

VI. Fill in the blanks :

*(i)* Two events are said to be equally likely if ......

*(ii)* A set of events is said to be independent if ......

*(iii)* If $P(A) \,.\, P(B) \,.\, P(C) = P(A \cap B \cap C)$, then the events $A, B, C$ are ......

*(iv)* Two events $A$ and $B$ are mutually exclusive if $P(A \cap B) = \ldots$ and are independent if $P(A \cap B) = \ldots$

*(v)* The probability of getting a multiple of 2 in a throw of a dice is 1/2 and of getting a multiple of 3 is 1/3. Hence probability of getting a multiple of 2 or 3 is ......

*(vi)* Let $A$ and $B$ be independent events and suppose the event $C$ has probability 0 or 1. Then $A, B$ and $C$ are ...... events.

*(vii)* If $A, B, C$ are pairwise independent and $A$ is independent of $B \cup C$, then $A, B, C$ are ...... independent.

*(viii)* A man has tossed 2 fair dice. The conditional probability that he has tossed two sixes, given that he has tossed at least one six is ......

*(ix)* Let $A$ and $B$ be two events such that $P(A) = 0{\cdot}3$ and $P(A \cup B) = 0{\cdot}8$. If $A$ and $B$ are independent events then $P(B) = \ldots$

**VII.** Each of following statements is either true or false. If it is true prove it, otherwise, give a counter example to show that it is false.

*(i)* The probability of occurrence of at least one of two events is the sum of the probability of each of the two events.

*(ii)* Mutually exclusive events are independent.

*(iii)* For any two events $A$ and $B$, $P(A \cap B)$ cannot be less than either $P(A)$ or $P(B)$.

*(iv)* The conditional probability of $A$ given $B$ is always greater than $P(A)$.

*(v)* If the occurrence of an event $A$ implies the occurrence of another event $B$ then $P(A)$ cannot exceed $P(B)$.

*(vi)* For any two events $A$ and $B$, $P(A \cup B)$ cannot be greater then either $P(A)$ or $P(B)$.

*(vii)* Mutually exclusive events are not independent.

*(viii)* Pairwise independence does not necessarily imply mutual independence.

*(ix)* Let $A$ and $B$ be events neither of which has probability zero. Then if $A$ and $B$ are disjoint, $A$ and $B$ are independent.

*(x)* The probability of any event is always a proper fraction.

*(xi)* If $0 < P(B) < 1$ so that $P(A \mid B)$ and $P(A \mid \overline{B})$ are both defined, then $P(A) = P(B)\,P(A \mid B) + P(\overline{B})\,P(A \mid \overline{B})$.

*(xii)* For two events $A$ and $B$ if

$P(A) = P(A \mid B) = 1/4$ and $P(A \mid \overline{B}) = 1/2$, then

*(a)* $A$ and $B$ are mutually exclusive.

*(b)* $A$ and $B$ are independent.

*(c)* $A$ is a sub-event of $B$.

*(d)* $P(\overline{A} \mid B) = 3/4$. **[Delhi Univ. B.Sc.(Stat. Hons.), 1992]**

*(xiii)* Two events can be independent and mutually exclusive simultaneously.

*(xiv)* Let $A$ and $B$ be events, neither of which has probability zero. Prove or disprove the following :

*(a)* If $A$ and $B$ are disjoint, $A$ and $B$ are independent.

*(b)* If $A$ and $B$ are independent, $A$ and $B$ are disjoint.

*(xv)* If $P(A) = 0$, then $A = \phi$.

# *Random Variables — Distribution Functions*

**5·1. Random Variable.** Intuitively by a *random variable (r.v)* we mean a real number $X$ connected with the outcome of a random experiment $E$. For example, if $E$ consists of two tosses of a coin, we may consider the random variable which is the number of heads ( 0, 1 or 2).

| | | | | |
|---|---|---|---|---|
| *Outcome :* | *HH* | *HT* | *TH* | *TT* |
| *Value of X :* | *2* | *1* | *1* | *0* |

Thus to each outcome $\omega$, there corresponds a real number $X(\omega)$. Since the points of the sample space $S$ correspond to outcomes, this means that a real number, which we denote by $X(\omega)$, is defined for each $\omega \in S$. From this standpoint, we define random variable to be a real function on S as follows:

*" Let S be the sample space associated with a given random experiment. A real-valued function defined on S and taking values in R $(-\infty, \infty)$ is called a one-dimensional random variable. If the function values are ordered pairs of real numbers (i.e., vectors in two-space) the function is said to be a two- dimensional random variable. More generally, an n-dimensional random variable is simply a function whose domain is S and whose range is a collection of n-tuples of real numbers (vectors in n- space)."*

For a mathematical and rigorous definition of the random variable, let us consider the probability space, the triplet (**S**, **B**, **P**), where $S$ is the sample space, *viz.*, space of outcomes, **B** is the $\sigma$-field of subsets in $S$, and **P** is a probability function on **B**.

Def. A random variable (r.v.) is a function $X(\omega)$ with domain $S$ and range $(-\infty, \infty)$ such that for every real number $a$, the event $[\omega : X(\omega) \leq a] \in \mathbf{B}$.

**Remarks: 1.** The refinement above is the same as saying that the function $X(\omega)$ is *measurable* real function on $(S, \mathbf{B})$.

2. We shall need to make probability statements about a random variable $X$ such as $P\{X \leq a\}$. For the simple example given above we should write $P\{X \leq 1\} = P\{HH, HT, TH\} = 3/4$. That is, $P(X \leq a)$ is simply the probability of the set of outcomes $\omega$ for which $X(\omega) \leq a$ or

$$P(X \leq a) = P\{\omega : X(\omega) \leq a\}$$

Since **P** is a measure on $(S, \mathbf{B})$ *i.e.*, **P** is defined on subsets of **B**, the above probability will be defined *only* if $\{\omega : X(\omega) \leq a\} \in \mathbf{B}$, which implies that $X(\omega)$ is a measurable function on $(S, \mathbf{B})$.

3. One-dimensional random variables will be denoted by capital letters, $X, Y, Z, \ldots$ etc. A typical outcome of the experiment (*i.e.*, a typical element of the sample space) will be denoted by $\omega$ or $e$. Thus $X(\omega)$ represents the real number which the random variable $X$ associates with the outcome $\omega$. The values which $X, Y, Z, \ldots$ etc., can assume are denoted by lower case letters *viz.*, $x, y, z, \ldots$ etc.

4. *Notations*. If $x$ is a real number, the set of all $\omega$ in $S$ such that $X(\omega) = x$ is denoted briefly by writing $X = x$. Thus

$$P(X = x) = P\{\omega : X(\omega) = x\}$$

Similarly $$P(X \leq a] = P\{\omega : X(\omega) \in [-\infty, a]\}$$

and $$P[a < X \leq b) = P(\omega : X(\omega) \in (a, b])$$

Analogous meanings are given to

$$P(X = a \text{ or } X = b) = P\{(X = a) \cup (X = b)\},$$

$$P(X = a \text{ and } X = b) = P\{(X = a) \cap (X = b)\}, \quad \text{etc.}$$

**Illustrations :** 1. If a coin is tossed, then

$$S = \{\omega_1, \omega_2\} \text{ where } \omega_1 = H, \ \omega_2 = T$$

$$X(\omega) = \begin{cases} 1, & \text{if } \omega = H \\ 0, & \text{if } \omega = T \end{cases}$$

$X(\omega)$ is a Bernoulli random variable. Here $X(\omega)$ takes only two values. A random variable which takes only a finite number of values is called *single*.

2. An experiment consists of rolling a die and reading the number of points on the upturned face. The most natural random variable $X$ to consider is

$$X(\omega) = \omega \ ; \ \omega = 1, 2, \ldots, 6$$

If we are interested in whether the number of points is even or odd, we consider a random variable $Y$ defined as follows :

$$Y(\omega) = \begin{cases} 0, & \text{if } \omega \text{ is even} \\ 1, & \text{if } \omega \text{ is odd} \end{cases}$$

3. If a dart is thrown at a circular target, the sample space $S$ is the set of all points $w$ on the target. By imagining a coordinate system placed on the target with the origin at the centre, we can assign various random variables to this experiment. A natural one is the two dimensional random variable which assigns to the point $\omega$, its rectangular coordinates $(x,y)$. Another is that which assigns $\omega$ its polar coordinates $(r, \theta)$. A one dimensional random variable assigns to each $\omega$ only one of the coordinates $x$ or $y$ (for cartesian system), $r$ or $\theta$ (for polar system). The event **E**, "that the dart will land in the first quadrant" can be described by a random variable which assigns to each point $w$ its polar coordinate $\theta$ so that $X(\omega) = \theta$ and then $\mathbf{E} = \{\omega : 0 \leq X(\omega) \leq \pi/2\}$.

4. If a pair of fair dice is tossed then $S = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$ and $n(S) = 36$. Let $X$ be a random variable with image set

$$X(S) = \{1,2,3,4,5,6\}$$

$$P(X = 1) = P\{1,1\} = 1/36$$

$$P(X = 2) = P\{(2,1),(2,2),(1,2)\} = 3/36$$

$$P(X = 3) = P\{(3,1),(3,2),(3,3),(2,3),(1,3)\} = 5/36$$

$P(X = 4) = P[(4,1), (4,2), (4,3), (4,4), (3,4), (2,4), (1,4)] = 7/36$

Similarly $P(X = 5) = 9/36$ and $P(X = 6) = 11/36$

**Some theorems on Random Variables.** Here we shall state (without proof) some of the fundamental results and theorems on random variables.

**Theorem 5·1.** A function $X(\omega)$ from $S$ to $R(-\infty, \infty)$ is a random variable if and only if

$$\{\omega : X(\omega) < a\} \in \mathbf{B}$$

**Theorem 5·2.** If $X_1$ and $X_2$ are random variables and $C$ is a constant then $CX_1, X_1 + X_2, X_1X_2$ are also random variables.

**Remark.** It will follow that $C_1X_1 + C_2X_2$ is a random variable for constants $C_1$ and $C_2$. In particular $X_1 - X_2$ is a r.v.

**Theorem 5·3.** If $\{X_n(\omega), n \geq 1\}$ are random variables then

$\sup_n X_n(\omega)$, $\inf_n X_n(\omega)$, $\lim\sup_{n\to\infty} X_n(\omega)$ and $\lim\inf_{n\to\infty} X_n(\omega)$ are all random variables, whenever they are finite for all $\omega$.

**Theorem 5·4.** If $X$ is a random variable then

(i) $\dfrac{1}{X}$ where $\left(\dfrac{1}{X}\right)(\omega) = \infty$ if $X(\omega) = 0$

(ii) $X_+(\omega) = \max[0, X(\omega)]$

(iii) $X_-(\omega) = -\min[0, X(\omega)]$

(iv) $|X|$

are random variables.

**Theorem 5·5.** If $X_1$ and $X_2$ are random variables then
(i) $\max[X_1, X_2]$ and (ii) $\min[X_1, X_2]$ are also random variables.

**Theorem 5·6.** If $X$ is a r.v. and $f(\cdot)$ is a continuous function, then $f(X)$ is a r.v.

**Theorem 5·7.** If $X$ is a r.v. and $f(\cdot)$ is an increasing function, then *$f(X)$ is a r.v.*

**Corollary.** If $f$ is a function of bounded variations on every finite interval $[a,b]$, and $X$ is a r.v. then $f(X)$ is a r.v.

*(proofs of the above theorems are beyond the scope of this book)*

## EXERCISE 5 (a)

1. Let $X$ be a one dimensional random variable. *(i)* If $a < b$, show that the two events $a < X \leq b$ and $X \leq a$ are disjoint, *(ii)* Determine the union of the two events in part *(i)*, *(iii)* show that $P(a < X \leq b) = P(X \leq b) - P(X \leq a)$.

2. Let a sample space $S$ consist of three elements $\omega_1$, $\omega_2$, and $\omega_3$. Let $P(\omega_1) = 1/4$, $P(\omega_2) = 1/2$ and $P(\omega_3) = 1/4$. If $X$ is a random variable defined on $S$ by $X(\omega_1) = 10$, $X(\omega_2) = -3$, $X(\omega_3) = 15$, find $P(-2 \leq X \leq 2)$.

3. Let $S = (e_1, e_2, ..., e_n)$ be the sample space of some experiment and let $E \subseteq S$ be some event associated with the experiment.

Define $\psi_E$, the *characteristic random variable* of $E$ as follows :

$$\psi_E(e_i) = \begin{cases} 1 & \text{if } e_i \in E \\ 0 & \text{if } e_i \notin E. \end{cases}$$

In other words, $\psi_E$ is equal to 1 if $E$ occurs, and $\psi_E$ is equal to 0 if $E$ does not occur.

Verify the following properties of characteristic random variables :

*(i)* $\psi_\varphi$ *is identically zero , i.e.,* $\psi_\varphi(e_i) = 0 \; ; \; i = 1, 2, \ldots, n$

*(ii)* $\psi_S$ *is identically one , i.e.,* $\psi_S(e_i) = 1 \; ; \; i = 1, 2, \ldots, n$

*(iii)* $E = F \Rightarrow \psi_E(e_i) = \psi_F(e_i) \; ; \; i = 1, 2, \ldots, n$ and *conversely*

*(iv) If* $E \subseteq F$ *then* $\psi_E(e_i) \leq \psi_F(e_i) \; ; \; i = 1, 2, \ldots, n$

*(v)* $\psi_E(e_i) + \psi_{\bar{E}}(e_i)$ *is identically* $1 \; : \; i = 1, 2, \ldots, n$

*(vi)* $\psi_{E \cap F}(e_i) = \psi_E(e_i)\, \psi_F(e_i) \; ; \; i = 1, 2, \ldots, n$

*(vii)* $\psi_{E \cup F}(e_i) = \psi_E(e_i) + \psi_F(e_i) - \psi_E(e_i)\, \psi_F(e_i)$, for $i = 1, 2, \ldots, n$.

**5.2. Distribution Function.** Let $X$ be a r.v. on ( S,B,P). Then the function :

$$F_X(x) = P(X \leq x) = P\{\omega : X(\omega) \leq x\}, \quad -\infty < x < \infty$$

is called the distribution function (d.f.) of $X$.

If clarity permits, we may write $F(x)$ instead of $F_X(x)$. ...(5·1)

**5·2·1. Properties of Distribution Function.** We now proceed to derive a number of properties common to all distribution functions.

**Property 1.** *If F is the d.f. of the r.v. X and if a < b, then*

$$P(a < X \leq b) = F(b) - F(a)$$

**Proof.** The events '$a<X\leq b$' and '$X\leq a$' are disjoint and their union is the event '$X\leq b$'. Hence by addition theorem of probability

$$P(a < X \leq b) + P(X \leq a) = P(X \leq b)$$

$$\Rightarrow \quad P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a) \qquad \text{...(5·2)}$$

**Cor. 1.**

$$\begin{aligned} P(a \leq X \leq b) &= P\{(X = a) \cup (a < X \leq b)\} \\ &= P(X = a) + P(a < X \leq b) \end{aligned}$$

(using additive property of $P$)

$$= P(X = a) + [F(b) - F(a)] \qquad \text{...(5·2 a)}$$

Similarly, we get

$$\begin{aligned} P(a < X < b) &= P(a < X \leq b) - P(X = b) \\ &= F(b) - F(a) - P(X = b) \qquad \text{...(5·2 b)} \end{aligned}$$

$$P(a \leq X < b) = P(a < X < b) + P(X = a)$$

$$= F(b) - F(a) - P(X = b) + P(X = a) \quad \text{...(5·2 c)}$$

**Remark.** When $P(X = a) = 0$ and $P(X = b) = 0$, all four events $a \le X \le b$, $a < X < b$, $a \le X < b$ and $a < X \le b$ have the same probability $F(b) - F(a)$.

**Property 2.** *If $F$ is the d.f. of one-dimensional r.v. $X$, then (i) $0 \le F(x) \le 1$, (ii) $F(x) \le F(y)$ if $x < y$.*

*In other words,* all distribution functions are monotonically non-decreasing and lie between 0 and 1.

**Proof.** Using the axioms of certainty and non-negativity for the probability function **P**, part *(i)* follows triviality from the definition of $F(x)$.

For part *(ii)*, we have for $x < y$,

$$F(y) - F(x) = P(x < X \le y) \ge 0 \quad \text{(Property 1)}$$

$$\Rightarrow \quad F(y) \ge F(x)$$

$$\Rightarrow \quad F(x) \le F(y) \text{ when } x < y \quad \text{...(5·3)}$$

**Property 3.** *If $F$ is d.f. of one-dimensional r.v. $X$, then*

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0$$

*and*

$$F(\infty) = \lim_{x \to \infty} F(x) = 1$$

**Proof.** Let us express the whole sample space $S$ as a countable union of disjoint events as follows :

$$S = \left[\bigcup_{n=1}^{\infty} (-n < X \le -n+1)\right] \cup \left[\bigcup_{n=0}^{\infty} (n < X \le n+1)\right]$$

$$\Rightarrow P(S) = \sum_{n=1}^{\infty} P(-n < X \le -n+1) + \sum_{n=0}^{\infty} P(n < X \le n+1)$$

$$(\because \mathbf{P} \text{ is additive})$$

$$\Rightarrow \quad 1 = \lim_{a \to \infty} \sum_{n=1}^{a} [F(-n+1) - F(-n)]$$

$$+ \lim_{b \to \infty} \sum_{n=0}^{b} [F(n+1) - F(n)]$$

$$= \lim_{a \to \infty} [F(0) - F(-a)] + \lim_{b \to \infty} [F(b+1) - F(0)]$$

$$= [F(0) - F(-\infty)] + [F(\infty) - F(0)]$$

$$\therefore \quad 1 = F(\infty) - F(-\infty) \quad \text{...(*)}$$

Since $-\infty < \infty$, $F(-\infty) \le F(\infty)$. Also

$$F(-\infty) \ge 0 \text{ and } F(\infty) \le 1 \quad \text{(Property 2)}$$

$$\therefore \qquad 0 \le F(-\infty) \le F(\infty) \le 1 \qquad (**)$$

(*) and (**) give $F(-\infty) = 0$ and $F(\infty) = 1$.

**Remarks. 1.** Discontinuities of $F(x)$ are at most countable.

**2.** $$F(a) - F(a-0) = \lim_{h \to 0} P(a-h \le X \le a),\ h > 0$$

$$\therefore \quad F(a) - F(a-0) = P(X = a)$$

and $$F(a+0) - F(a) = \lim_{h \to 0} P(a \le X \le a+h) = 0,\ h > 0$$

$$\Rightarrow \quad F(a+0) = F(a)$$

**5·3. Discrete Random Variable.** If a random variable takes at most a countable number of values, it is called a discrete random variable. *In other words, a real valued function defined on a discrete sample space is called a discrete random variable.*

**5·3·1. Probability Mass Function** (*and probability distribution of a discrete random variable*).

Suppose $X$ is a one-dimensional discrete random variable taking at most a countably infinite number of values $x_1, x_2, \ldots$ With each possible outcome $x_i$, we associate a number $p_i = P(X = x_i) = p(x_i)$, called the probability of $x_i$. The numbers $p(x_i)$; $i = 1, 2, \ldots$ must satisfy the following conditions:

$$(i) \quad p(x_i) \ge 0 \ \forall\ i, \quad (ii) \quad \sum_{i=1}^{\infty} p(x_i) = 1$$

This function $p$ is called the probability mass function of the random variable $X$ and the set $\{x_i, p(x_i)\}$ is called the probability distribution (p.d.) of the r.v. $X$.

**Remarks: 1.** The set of values which $X$ takes is called the *spectrum* of the random variable.

**2.** For discrete random variable, a knowledge of the probability mass function enables us to compute probabilities of arbitrary events. In fact, if $E$ is a set of real numbers, we have

$$P(X \in E) = \sum_{x \in E \cap S} p(x), \text{ where } S \text{ is the sample space.}$$

**Illustration.** Toss of coin, $S = \{H, T\}$. Let $X$ be the random variable defined by

$$X(H) = 1, \text{ i.e., } X = 1, \text{ if 'Head' occurs.}$$
$$X(T) = 0, \text{ i.e., } X = 0, \text{ if 'Tail' occurs.}$$

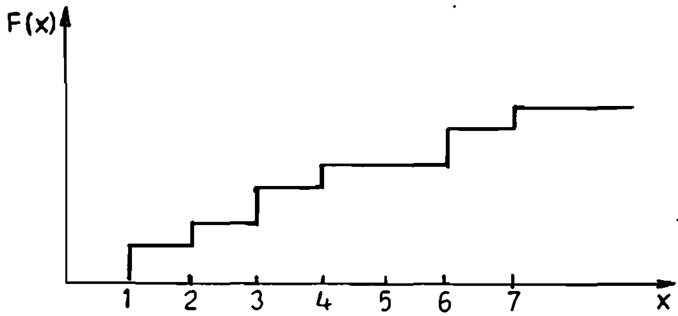
If the coin is 'fair' the probability function is given by

$$P(\{H\}) = P(\{T\}) = \tfrac{1}{2}$$

and we can speak of the probability distribution of the random variable $X$ as

$$P(X = 1) = P(\{H\}) = \tfrac{1}{2},$$
$$P(X = 0) = P(\{T\}) = \tfrac{1}{2},$$

**5·3·2. Discrete Distribution Function.** In this case there are a countable number of points $x_1, x_2, x_3, \ldots$ and numbers $p_i \geq 0$, $\sum_1^\infty p_i = 1$ such that $F(X) = \sum_{(i : x_i \leq x)} p_i$. For example if $x_i$ is just the integer $i$, $F(x)$ is a "*step function*" having jump $p_i$ at $i$, and being constant between each pair of integers.




**Theorem 5·5.** $p(x_j) = P(X = x_j) = F(x_j) - F(x_{j-1})$, where $F$ is the *d.f.* of $X$.

**Proof.** Let $x_1 < x_2 < \ldots$ We have

$$F(x_j) = P(X \leq x_j)$$

$$= \sum_{i=1}^{j} P(X = x_i) = \sum_{i=1}^{j} p(x_i)$$

and
$$F(x_{j-1}) = P(X \leq x_{j-1}) = \sum_{i=1}^{j-1} p(x_i)$$

$$\therefore \qquad F(x_j) - F(x_{j-1}) = p(x_j) \qquad \ldots(5\cdot5)$$

Thus, given the distribution function of discrete random variable, we can compute its probability mass function.

**Example 5·1.** *An experiment consists of three independent tosses of a fair coin. Let*

*X = The number of heads*

*Y = The number of head runs,*

*Z = The lenght of head runs,*

*a head run being defined as consecutive occurrence of at least two heads, its length then being the number of heads occurring together in three tosses of the coin.*

*Find the probability function of (i) X, (ii) Y, (iii) Z, (iv) X+Y and (v) XY and construct probability tables and draw their probability charts.*

**Solution.**

**Table 1**

| *S. No.* | *Elementary event* | *Random Variables* | | | | |
|---|---|---|---|---|---|---|
| | | $X$ | $Y$ | $Z$ | $X+Y$ | $XY$ |
| 1 | *HHH* | 3 | 1 | 3 | 4 | 3 |
| 2 | *HHT* | 2 | 1 | 2 | 3 | 2 |
| 3 | *HTH* | 2 | 0 | 0 | 2 | 0 |
| 4 | *HTT* | 1 | 0 | 0 | 1 | 0 |
| 5 | *THH* | 2 | 1 | 2 | 3 | 2 |
| 6 | *THT* | 1 | 0 | 0 | 1 | 0 |
| 7 | *TTH* | 1 | 0 | 0 | 1 | 0 |
| 8 | *TTT* | 0 | 0 | 0 | 0 | 0 |

Here sample space is.

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

(i) Obviously $X$ is a r.v. which can take the values 0, 1, 2, and 3

$$p(3) = P(HHH) = (1/2)^3 = 1/8$$

$$p(2) = P[HHT \cup HTH \cup THH]$$

$$= P(HHT) + P(HTH) + P(THH) = 1/8 + 1/8 + 1/8 = 3/8$$

Similarly $p(1) = 3/8$ and $p(0) = 1/8$.

These probabilities could also be obtained directly from the above table 1.

**Table 2**

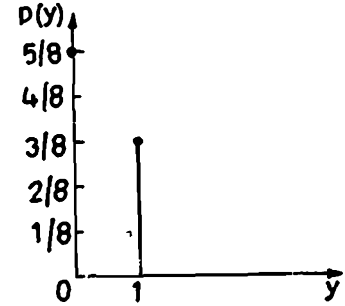
**Probability table of X**

| Values of $X$ $(x)$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p(x)$ | 1/8 | 3/8 | 3/8 | 1/8 |

**Table 3**

***(ii)* Probability Table of Y**

| Values of $Y$, $(y)$ | 0 1 |
|---|---|
| $p(y)$ | 5/8 3/8 |




Probability chart of $Y$

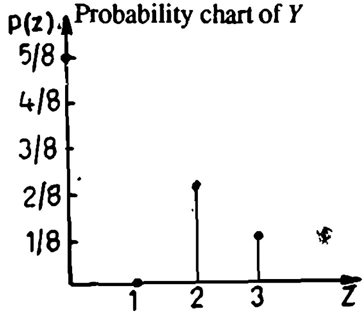
This is obvious from table 1.

*(iii)* From table 1 , we have

**Table 4**

**Probability Table of**

| Values of $Z$, $(z)$ | 0 1 2 3 |
|---|---|
| $p(z)$ | 5/8 0 2/8 1/8 |



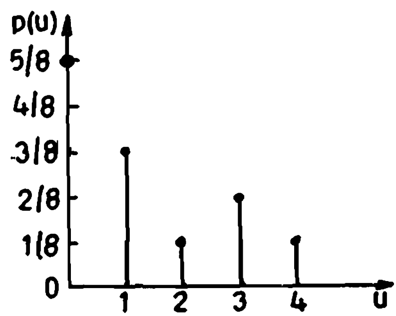

Probability chart of $Z$

*(iv)* Let $U = X + Y$. From table 1, we get

**Table 5**

**Probability Table of U**

| Values of $U$, $(u)$ | 0 1 2 3 4 |
|---|---|
| $p(u)$ | 1/8 3/8 1/8 2/8 1/8 |



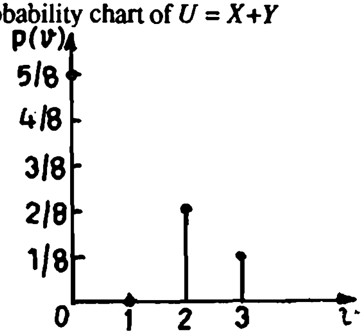

Probability chart of $U = X+Y$

*(v)* Let $V = XY$

**Table 6**

**Probability Table of V**

| Values of $V$, $(v)$ | 0 1 2 3 |
|---|---|
| $p(v)$ | 5/8 0 2/8 1/8 |




Probability chart of $V = XY$

**Example 5·2.** *A random variable X has the following probability distribution :*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| $p(x)$ : | 0 | $k$ | $2k$ | $2k$ | $3k$ | $k^2$ | $2k^2$ | $7k^2 + k$ |

*(i) Find k, (ii) Evaluate* $P(X < 6)$, $P(X \geq 6)$, *and* $P(0 < X < 5)$, *(iii) If* $P(X \leq c) > \frac{1}{2}$, *find the minimum value of c, and (iv) Determine the distribution function of X.* **[Madurai Univ. B.Sc., Oct. 1988]**

**Solution.** Since $\sum_{x=0}^{7} p(x) = 1$, we have

$\Rightarrow \quad k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$

$\Rightarrow \quad 10k^2 + 9k - 1 = 0$

$\Rightarrow \quad (10k - 1)(k + 1) = 0 \Rightarrow k = 1/10$

[$\because k = -1$, is rejected, since probability canot be negative.]

(*ii*) $P(X < 6) = P(X = 0) + P(X = 1) + \ldots + P(X = 5)$

$$= \frac{1}{10} + \frac{2}{10} + \frac{2}{10} + \frac{3}{10} + \frac{1}{100} = \frac{81}{100}$$

$$P(X \geq 6) = 1 - P(X < 6) = \frac{19}{100}$$

$P(0 < X < 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = 8k = 4/5$

(*iii*) $P(X \leq c) > \frac{1}{2}$. By trial, we get $c = 4$.

(*iv*)

| $X$ | $F_X(x) = P(X \leq x)$ |
|---|---|
| 0 | 0 |
| 1 | $k = 1/10$ |
| 2 | $3k = 3/10$ |
| 3 | $5k = 5/10$ |
| 4 | $8k = 4/5$ |
| 5 | $8k + k^2 = 81/100$ |
| 6 | $8k + 3k^2 = 83/100$ |
| 7 | $9k + 10k^2 = 1$ |

## EXERCISE 5 (b)

**1.** (*a*) A student is to match three historical events (Mahatma Gandhi's Birthday, India's freedom, and First World War) with three years (1947, 1914, 1896). If he guesses with no knowledge of the correct answers, what is the probability distribution of the number of answers he gets correctly ?

(*b*) From a lot of 10 items containing 3 defectives, a sample of 4 items is drawn at random. Let the random variable $X$ denote the number of defective items in the sample. Answer the following when the sample is drawn without replacement.

(*i*) Find the probability distribution of $X$,

(*ii*) Find $P(X \le 1)$, $P(X < 1)$ and $P(0 < X < 2)$

Ans. (*a*)

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p(x)$ | $\frac{1}{3}$ | $\frac{1}{2}$ | 0 | $\frac{1}{6}$ |

(*b*) (*i*)

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p(x)$ | $\frac{1}{6}$ | $\frac{1}{2}$ | $\frac{3}{10}$ | $\frac{1}{30}$ |

(*ii*) 2/3, 5/6, 1/2

2. (*a*) A random variable $X$ can take all non-negative integral values, and the probability that $X$ takes the value $r$ is proportional to $\alpha^r$ $(0 < \alpha < 1)$. Find $P(X = 0)$. **[Calcutta Univ. B.Sc. 1987]**

Ans. $P(X = r) = A\alpha^r$ ; $r = 0, 1, 2, \ldots$ ; $A = 1 - \alpha$ ; $P(X = 0) = A = 1 - \alpha$

(*b*) Suppose that the random variable $X$ has possible values 1, 2, 3, ... and $P(X = j) = 1/2^j$, $j = 1, 2, \ldots$ (*i*) Compute $P(X \text{ is even})$, (*ii*) Compute $P(X \ge 5)$, and (*iii*) Compute $P(X \text{ is divisible by } 3)$.

Ans. (*i*) 1/3, (*ii*) 1/16, and (*iii*) 1/7

3. (*a*) Let $X$ be a random variable such that

$$P(X = -2) = P(X = -1),\ P(X = 2) = P(X = 1) \text{ and}$$
$$P(X > 0) = P(X < 0) = P(X = 0).$$

Obtain the probability mass function of $X$ and its distribution function.

Ans.

| $X$ | : | −2 | −1 | 0 | 1 | 2 |
|---|---|---|---|---|---|---|
| $p(x)$ | : | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |
| $F(x)$ | : | $\frac{1}{6}$ | $\frac{2}{6}$ | $\frac{4}{6}$ | $\frac{5}{6}$ | 1 |

(*b*) A random variable $X$ assumes the values −3, −2, −1, 0, 1, 2, 3 such that

$$P(X = -3) = P(X = -2) = P(X = -1),$$
$$P(X = 1) = P(X = 2) = P(X = 3),$$
and $$P(X = 0) = P(X > 0) = P(X < 0),$$

Obtain the probability mass function of $X$ and its distribution function, and find further the probability mass function of $Y = 2X^2 + 3X + 4$.

**[Poona Univ. B.Sc., March 1991]**

Ans.

| $X$ | : | −3 | −2 | −1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|
| $p(x)$ | : | $\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{1}{3}$ | $\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{1}{9}$ |
| $Y$ | : | 13 | 6 | 3 | 4 | 9 | 18 | 31 |
| $p(y)$ | : | $\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{1}{3}$ | $\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{1}{9}$ |

4. *(a)* A random variable $X$ has the following probability function :

| Values of $X, x$ : | $-2$ | $-1$ | $0$ | $1$ | $2$ | $3$ |
|---|---|---|---|---|---|---|
| $p(x)$ : | $0{\cdot}1$ | $k$ | $0{\cdot}2$ | $2k$ | $0{\cdot}3$ | $k$ |

*(i)* Find the value of $k$, and calculate mean and variance.

*(ii)* Construct the c.d.f. $F(X)$ and draw its graph.

**Ans.** *(i)* 0·1, 0·8 and 2·16, *(ii)* $F(X) = 0{\cdot}1, 0{\cdot}2, 0{\cdot}4, 0{\cdot}6, 0{\cdot}9, 1{\cdot}0$

*(b)* Given the probability function

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p(x)$ | 0·1 | 0·3 | 0·5 | 0·1 |

Let $Y = X^2 + 2X$, then find *(i)* the probability function of $Y$, *(ii)* mean and variance of $Y$.

**Ans.** *(i)*

| $y$ | 0 | 3 | 8 | 15 |
|---|---|---|---|---|
| $p(y)$ | 0·1 | 0·3 | 0·5 | 0·1 |

, *(ii)* 6·4 ,16·24

5. A random variable $X$ has the following probability distribution :

| Values of $X$, $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| $p(x)$ | $a$ | $3a$ | $5a$ | $7a$ | $9a$ | $11a$ | $13a$ | $15a$ | $17a$ |

*(i)* Determine the value of $a$.

*(ii)* Find $P(X < 3)$, $P(X \geq 3)$, $P(0 < X < 5)$.

*(iii)* What is the smallest value of $x$ for which $P(X \leq x) > 0{\cdot}5$ ? and

*(iv)* Find out the distribution function of $X$ ?

**Ans.** *(i)* $a = 1/81$, *(ii)* 9/81, 72/81, 24/81, *(iii)* 6

| *(iv)* $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| $F(x)$ | $a$ | $4a$ | $9a$ | $16a$ | $25a$ | $36a$ | $49a$ | $64a$ | $81a$ |

6. *(a)* Let $p(x)$ be the probability function of a discrete random variable $X$ which assumes the values $x_1, x_2, x_3, x_4$, such that $2p(x_1) = 3p(x_2) = p(x_3) = 5p(x_4)$. Find probability distribution and cumulative probability distribution of $X$. **(Sardar Patel Univ. B.Sc. 1987)**

**Ans.**

| $x$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $p(x)$ | 15/16 | 10/16 | 30/16 | 6/16 |

*(b)* The following is the distribution function of a discrete random variable $X$ :

| $x$ | $-3$ | $-1$ | 0 | 1 | 2 | 3 | 5 | 8 |
|---|---|---|---|---|---|---|---|---|
| $f(x)$ | 0·10 | 0·30 | 0·45 | 0·5 | 0·75 | 0·90 | 0·95 | 1·00 |

*(i)* Find the probability distribution of $X$.

*(ii)* Find $P(X \text{ is even})$ and $P(1 \leq X \leq 8)$.

*(iii)* Find $P(X = -3 \mid X < 0)$ and $P(X \geq 3 \mid X > 0)$.

[ **Ans.** *(ii)* 0·30, 0·55, *(iii)* 1/3, 5/11].

7. If $$p(x) = \frac{x}{15};\ x = 1, 2, 3, 4, 5$$
$$= 0, \text{ elsewhere}$$

Find *(i)* $P\{X = 1 \text{ or } 2\}$, and *(ii)* $P\left\{\frac{1}{2} < X < \frac{5}{2} \mid X > 1\right\}$

**[Allahabad Univ. B.Sc., April 1992]**

**Hint.** $(i)\ P\{X = 1 \text{ or } 2\} = P(X = 1) + P(X = 2) = \frac{1}{15} + \frac{2}{15} = \frac{1}{5}$

$$(ii)\ P\left\{\frac{1}{2} < X < \frac{5}{2} \,\Big|\, X > 1\right\} = \frac{P\left\{\left(\frac{1}{2} < X < \frac{5}{2}\right) \cap X > 1\right\}}{P(X > 1)}$$

$$= \frac{P\{(X = 1 \text{ or } 2) \cap X > 1\}}{P(X > 1)} = \frac{P(X = 2)}{1 - P(X = 1)} = \frac{2/15}{1 - (1/15)} = \frac{1}{7}$$

8. The probability mass function of a random variable $X$ is zero except at the points $x = 0, 1, 2$. At these points it has the values $p(0) = 3c^3$, $p(1) = 4c - 10c^2$ and $p(2) = 5c - 1$ for some $c > 0$.

*(i)* Determine the value of $c$.

*(ii)* Compute the following probabilities, $P(X < 2)$ and $P(1 < X \leq 2)$.

*(iii)* Describe the distribution function and draw its graph.

*(iv)* Find the largest $x$ such that $F(x) < 1/2$.

*(v)* Find the smallest $x$ such that $F(x) \geq 1/3$. **[Poona Univ. B.Sc., 1987]**

Ans. *(i)* $\frac{1}{3}$, *(ii)* $\frac{1}{3}, \frac{2}{3}$, *(iv)* 1, *(v)* 1.

9. *(a)* Suppose that the random variable $X$ assumes three values 0,1 and 2 with probabilities $\frac{1}{3}$, $\frac{1}{6}$ and $\frac{1}{2}$ respectively. Obtain the distribution function of $X$..

**[Gujarat Univ. B.Sc., 1992]**

*(b)* Given that $f(x) = k(1/2)^x$ is a probability distribution for a random variable which can take on the values $x = 0, 1, 2, 3, 4, 5, 6$, find $k$ and find an expression for the corresponding cumulative probabilities $F(x)$.

**[Nagpur Univ. B.Sc., 1987)**

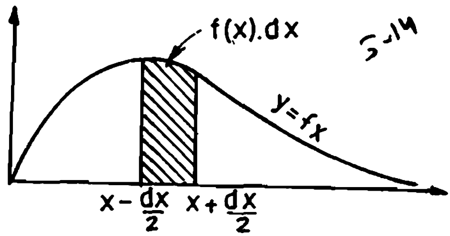
**5·4. Continuous Random Variable.** A random variable $X$ is said to be continuous if it can take all possible values between certain limits. *In other words, a random variable is said to be continuous when its different values cannot be put in 1-1 correspondence with a set of positive integers.*

A continuous random variable is a random variable that (at least conceptually) can be measured to any desired degree of accuracy. Examples of continuous random variables are age, height, weight etc.

**5·4·1. Probability Density Function** *(Concept and Definition).* Consider the small interval $(x, x + dx)$ of length $dx$ round the point $x$. Let $f(x)$ be any continuous function of $x$ so that $f(x)\,dx$ represents the probability that $X$ falls in the infinitesimal interval $(x, x+dx)$. Symbolically

$$P(x \le X \le x+dx) = f_X(x)\,dx \qquad \dots (5\cdot5)$$

In the figure, $f(x)\,dx$ represents the area bounded by the curve $y=f(x)$, $x$–axis and the ordinates at the points $x$ and $x+dx$. The function $f_X(x)$ so defined is known as *probability density function or simply density function of random variable X and is usually abbreviated as p.d.f.* The expression, $f(x)\,dx$, usually written as $dF(x)$, is known as the *probability differential* and the curve $y=f(x)$ is known as the *probability density curve* or simply *probability curve.*




**Definition.** p.d.f. $f_X(x)$ of the *r.v.* $X$ is defined as :

$$f_X(x) = \lim_{\delta x \to 0} \frac{P(x \le X \le x+\delta x)}{\delta x} \qquad \dots (5\cdot5\,a)$$

The probability for a variate value to lie in the interval $dx$ is $f(x)\,dx$ and hence the probability for a variate value to fall in the finite interval $[\alpha, \beta]$ is :

$$P(\alpha \le X \le \beta) = \int_\alpha^\beta f(x)\,dx \qquad \dots (5\cdot5\,b)$$

which represents the area between the curve $y=f(x)$, $x$–axis and the ordinates at $x=\alpha$ and $x=\beta$. Further since total probability is unity, we have $\int_a^b f(x)\,dx = 1$, where $[a, b]$ is the range of the random variable $X$. The range of the variable may be finite or infinite.

The probability density function (*p.d.f.*) of a random variable (*r.v.*) $X$ usually denoted by $f_X(x)$ or simply by $f(x)$ has the following obvious properties

*(i)* $f(x) \ge 0, \; -\infty < x < \infty$ $\qquad \dots (5\cdot5\,c)$

*(ii)* $\int_{-\infty}^{\infty} f(x)\,dx = 1$ $\qquad \dots (5\cdot5\,d)$

*(iii)* The probability $P(E)$ given by

$$P(E) = \int_E f(x)\,dx \qquad \dots (5\cdot5\,e)$$

is well defined for any event $E$.

**Important Remark.** In case of discrete random variable, the probability at a point, *i.e.*, $P(x=c)$ is not zero for some fixed $c$. However, in case of continuous random variables the probability at a point is always zero, *i.e.*, $P(x=c)=0$ for all possible values of $c$. This follows directly from (5·5 *b*) by taking $\alpha = \beta = c$.

This also agrees with our discussion earlier that $P(E) = 0$ does not imply that the event $E$ is null or impossible event. This property of continuous r.v., *viz.*,

$$P(X = c) = 0, \ \forall \ c \quad \ldots (5\cdot5 f)$$

leads us to the following important result :

$$P(\alpha \leq X \leq \beta) = P(\alpha \leq X < \beta) = P(\alpha < X \leq \beta) = P(\alpha < X < \beta) \quad \ldots (5\cdot5 g)$$

*i.e.*, in case of continuous r.v., it does matter whether we include the end points of the interval from $\alpha$ to $\beta$.

However, this result is in general not true for discrete random variables.

**5·4·2. Various Measures of Central Tendency, Dispersion, Skewness, and Kurtosis for Continuous Probability Distribution.** The formulae for these measures in case of discrete frequency distribution can be easily extended to the case of continuous probability distribution by simply replacing $p_i = f_i/N$ by $f(x)\,dx, x_i$ by $x$ and the summation over '$i$' by integration over the specified range of the variable $X$.

Let $f_X(x)$ *or* $f(x)$ be the *p.d.f.* of a random variable $X$ where $X$ is defined from $a$ to $b$. Then

*(i)* *Arithmetic mean* $= \int_a^b x\, f(x)\, dx$ ...(5·6)

*(ii) Harmonic mean.* Harmonic mean $H$ is given by

$$\frac{1}{H} = \int_a^b \left(\frac{1}{x}\right) f(x)\, dx \quad \ldots(5\cdot6\ a)$$

*(iii) Geometric mean.* Geometric mean $G$ is given by

$$\log G = \int_a^b \log x\, f(x)\, dx \quad \ldots(5\cdot6\ b)$$

*(iv)* $\mu_r'$ (about origin) $= \int_a^b x^r f(x)\, dx$ ...(5·7)

$\mu_r'$ (about the point $x = A$) $= \int_a^b (x - A)^r f(x)\, dx$ ...(5·7 *a*)

and $\mu_r$ (about mean) $= \int_a^b (x - mean)^r f(x)\, dx$ ...(5·7 *b*)

In particular, from (5·7), we have

$\mu_1'$ (about origin) $=$ *Mean* $= \int_a^b x\, f(x)\, dx$

and $\mu_2' = \int_a^b x^2 f(x)\, dx$

Hence $\mu_2 = \mu_2' - \mu_1'^2 = \int_a^b x^2 f(x)\, dx - \left(\int_a^b x f(x)\, dx\right)^2$ ... (5·7 *c*)

From (5·7), on putting $r=3$ and 4 respectively, we get the values of $\mu_3'$ and $\mu_4'$ and consequently the moments about mean can be obtained by using the relations :

and
$$\left.\begin{aligned}\mu_3 &= \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 \\ \mu_4 &= \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4\end{aligned}\right\} \quad \ldots (5.7\ d)$$

and hence $\beta_1$ and $\beta_2$ can be computed.

*(v) Median.* Median is the point which divides the entire distribution in two equal parts. In case of continuous distribution, median is the point which divides the total area into two equal parts. Thus if $M$ is the median, then

$$\int_a^M f(x)\,dx = \int_M^b f(x)\,dx = \frac{1}{2} \quad \ldots (5.8)$$

Thus solving

$$\int_a^M f(x)\,dx = \frac{1}{2} \quad or \quad \int_M^b f(x)\,dx = \frac{1}{2} \quad \ldots (5.8\ a)$$

for $M$, we get the value of median.

*(vi) Mean Deviation.* Mean deviation about the mean $\mu_1'$ is given by

$$M.D. = \int_a^b |\ x - mean\ |\ f(x)\,dx \quad \ldots (5.9)$$

*(vii) Quartiles and Deciles.* $Q_1$ and $Q_3$ are given by the equations

$$\int_a^{Q_1} f(x)\,dx = \frac{1}{4} \quad \text{and} \quad \int_a^{Q_3} f(x)\,dx = \frac{3}{4} \quad \ldots (5.10)$$

$D_i$, $i$ th decile is given by

$$\int_a^{D_i} f(x)\,dx = \frac{i}{10} \quad \ldots (5.10\ a)$$

*(viii) Mode.* Mode is the value of $x$ for which $f(x)$ is maximum. Mode is thus the solution of

$$f'(x) = 0 \text{ and } f''(x) < 0 \quad \ldots (5.11)$$

provided it lies in $[a,b]$.

**Example 5.3.** *The diameter of an electric cable, say X, is assumed to be a continuous random variable with p.d.f.* $f(x) = 6x(1-x)$, $0 \le x \le 1$.

*(i) Check that above is p.d.f.,*

*(ii) Determine a number b such that* $P(X < b) = P(X > b)$

[Aligarh Univ. B.Sc.(Hons).1990]

**Solution.** Obviously, for $0 \le x \le 1$, $f(x) \ge 0$

Now
$$\int_0^1 f(x)\,dx = 6\int_0^1 x(1-x)\,dx$$
$$= 6\int_0^1 (x - x^2)\,dx = 6\left|\frac{x^2}{2} - \frac{x^3}{3}\right|_0^1 = 1$$

Hence $f(x)$ is the *p.d.f.* of *r.v.* $X$

*(ii)* $$P(X < b) = P(X > b) \quad \ldots(*)$$

$$\Rightarrow \qquad \int_0^b f(x)\,dx = \int_b^1 f(x)\,dx$$

$$\Rightarrow \qquad 6\int_0^b x(1-x)\,dx = 6\int_b^1 x(1-x)\,dx$$

$$\Rightarrow \qquad \left|\frac{x^2}{2}-\frac{x^3}{3}\right|_0^b = \left|\frac{x^2}{2}-\frac{x^3}{3}\right|_b^1$$

$$\Rightarrow \qquad \left(\frac{b^2}{2}-\frac{b^3}{3}\right) = \left[\left(\frac{1}{2}-\frac{1}{3}\right)-\left(\frac{b^2}{2}-\frac{b^3}{3}\right)\right]$$

$$\Rightarrow \qquad 3b^2-2b^3 = [\,1-3b^2+2b^3\,]$$

$$\Rightarrow \qquad 4b^3-6b^2+1=0$$

$$(2b-1)(2b^2-2b-1)=0$$

$$\Rightarrow \qquad 2b-1=0 \quad or \quad 2b^2-2b-1=0$$

Hence $b = 1/2$ is the only real value lying between 0 and 1 and satisfying (*).

**Example 5·4.** *A continuous random variable $X$ has a p.d.f.*

$$f(x) = 3x^2,\ 0 \le x \le 1. \quad \textit{Find a and b such that}$$

(i) $P\{X \le a\} = P\{X > a\}$, and

(ii) $P\{X > b\} = 0{\cdot}05$. **[Calicut Univ. B.Sc., Sept. 1988]**

**Solution.** *(i)* Since $P(X \le a) = P(X > a)$,
each must be equal to $1/2$, because total probability is always one.

$$\therefore \qquad P(X \le a) = \frac{1}{2} \quad \Rightarrow \quad \int_0^a f(x)\,dx = \frac{1}{2}$$

$$\Rightarrow \qquad 3\int_0^a x^2\,dx = \frac{1}{2} \quad \Rightarrow \quad 3\left|\frac{x^3}{3}\right|_0^a = \frac{1}{2}$$

$$\Rightarrow \qquad a^3 = \frac{1}{2} \quad \Rightarrow \quad a = \left(\frac{1}{2}\right)^{\frac{1}{3}}$$

$$(ii) \qquad P(X > b) = 0{\cdot}05 \quad \Rightarrow \quad \int_b^1 f(x)\,dx = 0{\cdot}05$$

$$\Rightarrow \qquad 3\left|\frac{x^3}{3}\right|_b^1 = \frac{1}{20} \quad \Rightarrow \quad 1-b^3 = \frac{1}{20}$$

$$\Rightarrow \qquad b^3 = \frac{19}{20} \quad \Rightarrow \quad b = \left(\frac{19}{20}\right)^{\frac{1}{3}}.$$

**Example 5·5.** *Let $X$ be a continuous random variate with p.d.f.*

$$\begin{aligned} f(x) &= ax, \quad 0 \le x \le 1\\ &= a, \quad 1 \le x \le 2\\ &= -ax+3a, \quad 2 \le x \le 3\\ &= 0, \textit{ elsewhere}\end{aligned}$$

*(i) Determine the constant a.*

*(ii) Compute $P(X \leq 1{\cdot}5)$.* **[Sardar Patel Univ. B.Sc., Nov.1988]**

**Solution.** *(i)* Constant '$a$' is determined from the consideration that total probability is unity, *i.e.*,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_0^1 f(x)\,dx + \int_1^2 f(x)\,dx + \int_2^3 f(x)\,dx + \int_3^{\infty} f(x)\,dx = 1$$

$$\Rightarrow \int_0^1 ax\,dx + \int_1^2 a.dx + \int_2^3 (-ax+3a)\,dx = 1$$

$$\Rightarrow a\left|\frac{x^2}{2}\right|_0^1 + a\left|x\right|_1^2 + a\left|-\frac{x^2}{2}+3x\right|_2^3 = 1$$

$$\Rightarrow \frac{a}{2} + a + a\left[\left(-\frac{9}{2}+9\right) - (-2+6)\right] = 1$$

$$\Rightarrow \frac{a}{2} + a + \frac{a}{2} = 1 \;\Rightarrow\; 2a = 1 \;\Rightarrow\; a = \frac{1}{2}$$

*(ii)* $P(X \leq 1{\cdot}5) = \int_{-\infty}^{1\cdot5} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_0^1 f(x)\,dx + \int_1^{1\cdot5} f(x)\,dx$

$$= a\int_0^1 x\,dx + \int_1^{1\cdot5} a.dx$$

$$= a\left|\frac{x^2}{2}\right|_0^1 + a\left|x\right|_1^{1\cdot5} = \frac{a}{2} + 0{\cdot}5\,a$$

$$= a = \frac{1}{2} \qquad \left[\because a = \frac{1}{2},\ \textit{Part (i)}\right]$$

**Example 5·6.** *A probability curve $y = f(x)$ has a range from 0 to $\infty$. If $f(x) = e^{-x}$, find the mean and variance and the third moment about mean.*

**[Andhra Univ. B.Sc. 1988; Delhi Univ. B.Sc. Sept. 1987]**

**Solution.**

$\mu_r$ (rth moment about origin) $= \int_0^{\infty} x^r f(x)\,dx$

$$= \int_0^{\infty} x^r e^{-x}\,dx = \Gamma(r+1) = r\,!$$

(Using Gamma Integral)

Substituting $r = 1, 2$ and $3$ successively, we get

$$\text{Mean} = \mu_1' = 1\,! = 1,\quad \mu_2' = 2\,! = 2,\quad \mu_3' = 3\,! = 6$$

Hence variance $= \mu_2 = \mu_2' - \mu_1'^2 = 2 - 1 = 1$

and $\quad \mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 = 6 - 3\times2 + 2 = 2$

**Example 5·7.** *In a continuous distribution whose relative frequency density is given by*

$$f(x) = y_o \,.\, x(2-x)\,,\ 0 \le x \le 2,$$

*find mean, variance,* $\beta_1$ *, and* $\beta_2$ *and hence show that the distribution is symmetrical. Also (i) find mean deviation about mean and (ii) show that for this distribution* $\mu_{2n+1} = 0$, *(iii) find the mode, harmonic mean and median.*

**[Delhi Univ. B.Sc.(Stat. Hons.), 1992; B.Sc., Oct. 1992]**

**Solution.** Since total probability is unity, we have

$$\int_0^2 f(x)\,dx = 1$$

$$\Rightarrow \qquad y_o \int_0^2 x(2-x)\,dx = 1 \quad \Rightarrow \quad y_o = 3/4$$

$$\therefore \qquad f(x) = \frac{3}{4} x(2-x)$$

$$\mu_r' = \int_0^2 x^r f(x)\,dx = \frac{3}{4}\int_0^2 x^{r+1}(2-x)\,dx = \frac{3 \cdot 2^{r+1}}{(r+2)(r+3)}$$

In particular

$$\text{Mean} = \mu_1' = \frac{3 \cdot 2^2}{3 \cdot 4} = 1, \quad \mu_2' = \frac{3 \cdot 2^3}{4 \cdot 5} = \frac{6}{5},$$

$$\mu_3' = \frac{3 \cdot 2^4}{5 \cdot 6} = \frac{8}{5}, \quad \text{and} \quad \mu_4' = \frac{3 \cdot 2^5}{6 \cdot 7} = \frac{16}{7}$$

$$\text{Hence varience} = \mu_2 = \mu_2' - \mu_1'^2 = \frac{6}{5} - 1 = \frac{1}{5}$$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 = \frac{8}{5} - 3 \,.\, \frac{6}{5} \,.\, 1 + 2 = 0$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4 = \frac{16}{7} - 4 \,.\, \frac{8}{5} \,.\, 1 + 6 \,.\, \frac{6}{5} \,.\, 1 - 3 \,.\, 1 = \frac{3}{35}$$

$$\therefore \qquad \beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0 \quad \text{and} \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3/35}{(1/5)^2} = \frac{15}{7}$$

Since $\beta_1 = 0$, the distributior is *symmetrical.*

*Mean deviation about mean*

$$= \int_0^2 |x-1|\, f(x)\,dx$$

$$= \int_0^1 |x-1|\, f(x)\,dx + \int_1^2 |x-1|\, f(x)\,dx$$

$$= \frac{3}{4}\left[\int_0^1 (1-x)\,x(2-x)\,dx + \int_1^2 (x-1)\,x(2-x)\,dx\right]$$

$$= \frac{3}{4}\left[\int_0^1 (2x - 3x^2 + x^3)\,dx + \int_1^2 (3x^2 - x^3 - 2x)\,dx\right]$$

$$= \frac{3}{4}\left[\left| x^2 - \frac{3 \cdot x^3}{3} + \frac{x^4}{4}\right|_0^1 + \left| 3 \cdot \frac{x^3}{3} - \frac{x^4}{4} - \frac{2x^2}{2}\right|_1^2\right] = \frac{3}{8}$$

$$\mu_{2n+1} = \int_0^2 (x - mean)^{2n+1} f(x)\,dx$$

$$= \frac{3}{4}\int_0^2 (x-1)^{2n+1} x(2-x)\,dx$$

$$= \frac{3}{4}\int_{-1}^1 t^{2n+1}(t+1)(1-t)\,dt \qquad (x - 1 = t)$$

$$= \frac{3}{4}\int_{-1}^1 t^{2n+1}(1-t^2)\,dt$$

Since $t^{2n+1}$ 1 is an odd function of $t$ and $(1 - t^2)$ is an even function of $t$, the integrand $t^{2n+1}(1 - t^2)$ is an odd function of $t$.

Hence $\mu_{2n+1} = 0.$

Now $f'(x) = \frac{3}{4}(2 - 2x) = 0 \quad \Rightarrow \quad x = 1$

and $f''(x) = \frac{3}{4}(-2) = -\frac{3}{2} < 0$

Hence mode $= 1$

Harmonic mean $H$ is given by

$$\frac{1}{H} = \int_0^2 \frac{1}{x} f(x)\,dx$$

$$= \frac{3}{4}\int_0^2 (2-x)\,dx = \frac{3}{2}$$

$$\Rightarrow \qquad H = \frac{2}{3}$$

If $M$ is the median, then

$$\int_0^M f(x)\,dx = \frac{1}{2}$$

$$\Rightarrow \qquad \frac{3}{4}\int_0^M x(2-x)\,dx = \frac{1}{2}$$

$$\Rightarrow \qquad \left| x^2 - \frac{x^3}{3}\right|_0^M = \frac{2}{3}$$

$$\Rightarrow \qquad 3M^2 - M^3 = 2$$

$$\Rightarrow \qquad M^3 - 3M^2 + 2 = 0$$

$$\Rightarrow \qquad (M-1)(M^2 - 2M - 2) = 0$$

The only value of $M$ lying in $[0, 2]$ is $M = 1$. Hence median is 1.

**Aliter.** Since we have proved that distribution is symmetrical,

$$\text{Mode} = \text{Median} = \text{Mean} = 1$$

**Example 5·8.** *The elementary probability law of a continuous random variable X is*

$$f(x) = y_o e^{-b(x-a)}, \quad a \le x < \infty, \quad b > 0$$

*where a, b and $y_o$ are constants.*

*Show that $y_o = b = 1/\sigma$ and $a = m - \sigma$, where m and $\sigma$ are respectively the mean and standard deviation of the distribution. Show also that $\beta_1 = 4$ and $\beta_2 = 9$.* **[Gauhati Univ. B.Sc., 1992]**

**Solution.** Since total probability is unity,

$$\int_a^\infty f(x)\,dx = 1 \quad \Rightarrow \quad y_o \int_a^\infty e^{-b(x-a)}\,dx = 1$$

$$\Rightarrow \qquad y_o \left| \frac{e^{-b(x-a)}}{-b} \right|_a^\infty = 1 \quad \Rightarrow \quad y_o \frac{1}{b} = 1, \ (b > 0)$$

$$\Rightarrow \qquad y_o = b$$

$\mu_r'$ ( $r$th moment about the point '$x = a$' )

$$= \int_a^\infty (x-a)^r f(x)\,dx = b \int_a^\infty (x-a)^r e^{-b(x-a)}\,dx$$

$$= b \int_0^\infty t^r e^{-bt}\,dt \qquad \text{[ On putting } x - a = t \text{ ]}$$

$$= b\, \frac{\Gamma(r+1)}{b^{r+1}} = \frac{r!}{b^r} \qquad \text{[ Using Gamma Integral ]}$$

In particular

$$\mu_1' = 1/b, \ \mu_2' = 2/b^2, \ \mu_3' = 6/b^3, \ \mu_4' = 24/b^4$$

$$\therefore \qquad m = Mean = a + \mu_1' = a + (1/b)$$

and $\qquad \sigma^2 = \mu_2 = \mu_2' - \mu_1'^2 = 1/b^2$

$$\Rightarrow \qquad \sigma = \frac{1}{b} \ \text{ and } \ m = a + \frac{1}{b} = a + \sigma$$

Hence $\qquad y_o = b = \dfrac{1}{\sigma}$ and $a = m - \sigma$

Also $\qquad \mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 = \dfrac{1}{b^3}(6 - 3\cdot2 + 2) = \dfrac{2}{b^3} = 2\sigma^3$

and $\qquad \mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$

$$= \frac{1}{b^4}(24 - 4.6.1 + 6.2.1 - 3) = \frac{9}{b^4} = 9\sigma^4$$

Hence $\beta_1 = \mu_3^2/\mu_2^3 = 4\sigma^6/\sigma^6 = 4$ and $\beta_2 = \mu_4/\mu_2^2 = 9\sigma^4/\sigma^4 = 9$

**Example 5·9.** *For the following probability distribution*

$$dF = y_o \cdot e^{-|x|}\, dx\,, \quad -\infty < x < \infty$$

*show that* $y_o = \frac{1}{2}$, $\mu_1' = 0$, $\sigma = \sqrt{2}$ *and mean deviation about mean = 1.*

**Solution.** We have $\int_{-\infty}^{\infty} f(x)\, dx = 1$

$$\Rightarrow \quad y_o \int_{-\infty}^{\infty} e^{-|x|}\, dx = 1 \quad \Rightarrow \quad 2y_o \int_0^{\infty} e^{-|x|}\, dx = 1,$$

(since $e^{-|x|}$ is an even function of $x$)

$$\Rightarrow \quad 2y_o \int_0^{\infty} e^{-x}\, dx = 1,$$ (since in $0 \leq x < \infty$, $|x| = x$)

$$\Rightarrow \quad 2y_o \left| \frac{e^{-x}}{-1} \right|_0^{\infty} = 1 \quad \Rightarrow \quad 2y_o = 1, \ \textit{i.e.,}\ y_o = \frac{1}{2}$$

$$\mu_1' \text{ (about origin)} = \int_{-\infty}^{\infty} x\, f(x)\, dx = \frac{1}{2}\int_{-\infty}^{\infty} x\, e^{-|x|}\, dx$$

$$= 0,$$

(since the integrand $x \cdot e^{-|x|}$ is an odd function of $x$)

$$\mu_2' = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \frac{1}{2}\int_{-\infty}^{\infty} x^2\, e^{-|x|}\, dx$$

$$= \frac{1}{2}\, 2 \int_0^{\infty} x^2\, e^{-|x|}\, dx$$

[since the integrand $x^2 e^{-|x|}$ is an even function of $x$]

$$\therefore \quad \mu_2' = \int_0^{\infty} x^2 e^{-x}\, dx = \Gamma(3)$$ (on using Gamma Integral)

$$\Rightarrow \quad \mu_2' = 2! = 2$$

Now $\quad \sigma^2 = \mu_2 = \mu_2' - \mu_1'^2 = 2$

$$\text{M.D. about mean} = \int_{-\infty}^{\infty} |x - mean|\, f(x)\, dx$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} |x|\, e^{-|x|}\, dx \qquad (\because \text{Mean} = \mu_1' = 0)$$

$$= \frac{1}{2} \cdot 2 \int_0^{\infty} |x|\, e^{-|x|}\, dx$$

$$= \int_0^{\infty} x\, e^{-x}\, dx = \Gamma(2) = 1$$

**Example 5·10.** *A random variable X has the probability law :*

$$dF(x) = \frac{x}{b^2} \cdot e^{-x^2/2b^2}\, dx, \quad 0 \le x < \infty$$

*Find the distance between the quartiles and show that the ratio of this distance to the standard devation of X is independent of the patameter 'b'.*

**Solution.** If $Q_1$ and $Q_3$ are the first and third quartiles respectively, we have

$$\int_0^{Q_1} f(x)\,dx = \frac{1}{4} \quad \Rightarrow \quad \frac{1}{b^2}\int_0^{Q_1} x\, e^{-x^2/2b^2}\, dx = \frac{1}{4}$$

Put $\quad y = \dfrac{x^2}{2b^2} \quad$ *then* $\quad dy = \dfrac{x}{b^2}\, dx$

$$\therefore \quad \int_0^{Q_1^2/2b^2} e^{-y}\, dy = \frac{1}{4} \quad \Rightarrow \quad \left|\frac{e^{-y}}{-1}\right|_0^{Q_1^2/2b^2} = \frac{1}{4}$$

$$\Rightarrow \quad 1 - e^{-Q_1^2/2b^2} = \frac{1}{4} \quad \Rightarrow \quad e^{-Q_1^2/2b^2} = \frac{3}{4}$$

$$\Rightarrow \quad Q_1 = \sqrt{2}\, b \sqrt{\log(4/3)}$$

Again we have $\int_0^{Q_3} f(x)\,dx = \dfrac{3}{4}$ which, on proceeding similarly, will give

$$1 - e^{-Q_3^2/2b^2} = 3/4 \quad \Rightarrow \quad e^{-Q_3^2/2b^2} = 1/4$$

$$\Rightarrow \quad Q_3 = \sqrt{2}\, b \sqrt{\log(4)}$$

The distance between the quartiles is given by

$$Q_3 - Q_1 = \sqrt{2}\, b\,[\sqrt{\log 4} - \sqrt{\log(4/3)}\,]$$

$$\mu_1' = \int_0^{\infty} x\, f(x)\,dx = \int_0^{\infty} x\, \frac{x}{b^2}\, e^{-x^2/2b^2}\, dx$$

$$= \int_0^{\infty} \sqrt{2}\, b y^{1/2}\, e^{-y}\, dy \qquad \left(y = \frac{x^2}{2b^2}\right)$$

$$= \sqrt{2}\, b \int_0^{\infty} e^{-y}\, y^{(3/2)-1}\, dy$$

$$= \sqrt{2}\, b\, \Gamma\left(\frac{3}{2}\right) = \sqrt{2}\,.b\,.\,\frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) = \sqrt{2}\, b \frac{\sqrt{\pi}}{2} = b\sqrt{(\pi/2)}$$

$$\mu_2' = \int_0^{\infty} x^2 f(x)\,dx = \int_0^{\infty} x^2\, \frac{x}{b^2}\, e^{-x^2/2b^2}\, dx$$

$$= 2b^2 \int_0^{\infty} y\, e^{-y}\, dy \qquad \left(y = \frac{x^2}{2b^2}\right)$$

$$= 2b^2\, \Gamma(2) = 2b^2 \,.\, 1\,! = 2b^2$$

$$\therefore \quad \sigma^2 = \mu_2 = \mu_2' - \mu_1'^2 = 2b^2 - b^2 \cdot \frac{\pi}{2} = b^2\left(2 - \frac{\pi}{2}\right)$$

$$\Rightarrow \quad \sigma = b\sqrt{2 - (\pi/2)}$$

Hence $\dfrac{Q_3 - Q_1}{\sigma} = \dfrac{\sqrt{2}\,[\sqrt{\log 4} - \sqrt{\log(4/3)}\,]}{\sqrt{2 - (\pi/2)}}$,

which is independent of the parameter '$b$'.

**Example 5·11.** *Prove that the geometric mean G of the distribution*

$$dF = 6(2-x)(x-1)\,dx, \quad 1 \le x \le 2$$

*is given by* $6 \log(16G) = 19$. **[Kanpur Univ. B.Sc., Oct. 1992]**

**Solution.** By definition, we have

$$\log G = \int_1^2 \log x\, f(x)\,dx = 6\int_1^2 \log x\,(2-x)(x-1)\,dx$$

$$= -6\int_1^2 (x^2 - 3x + 2)\log x\,dx$$

Integrating by parts, we get

$$\log G = -6\left[\left|\left(\frac{x^3}{3} - \frac{3x^2}{2} + 2x\right)\log x\right|_1^2 - \int_1^2 \left(\frac{x^3}{3} - \frac{3x^2}{2} + 2x\right)\frac{1}{x}\,dx\right]$$

$$= -4\log 2 + 6 \times \frac{19}{36} \qquad \text{(on simplification)}$$

$$\therefore \quad \log G + 4\log 2 = \frac{19}{6} \quad \Rightarrow \quad \log G + \log 2^4 = \frac{19}{6}$$

$$\Rightarrow \quad \log G + \log 16 = \frac{19}{6} \quad \Rightarrow \quad \log(16\,G) = \frac{19}{6}$$

$$\Rightarrow \quad 6\log(16\,G) = 19$$

**Example 5·12.** *The time one has to wait for a bus at a downtown bus stop is observed to be random phenomenon (X) with the following probability density function :*

$$\begin{aligned} f_X(x) &= 0, && \text{for } x < 0 \\ &= \tfrac{1}{9}(x+1), && \text{for } 0 \le x < 1 \\ &= \tfrac{4}{9}(x - \tfrac{1}{2}), && \text{for } 1 \le x < \tfrac{3}{2} \\ &= \tfrac{4}{9}(\tfrac{5}{2} - x), && \text{for } \tfrac{3}{2} \le x < 2 \\ &= \tfrac{1}{9}(4 - x), && \text{for } 2 \le x < 3 \\ &= \tfrac{1}{9}, && \text{for } 3 \le x < 6 \end{aligned}$$

$$= 0, \qquad \text{for} \quad 6 \le x,$$

*Let the events A and B be defined as follows :*

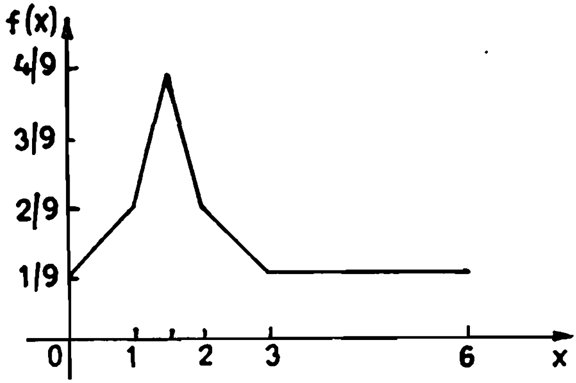
*A : One waits between 0 to 2 minutes inclusive:*

*B : One waits between 0 to 3 minutes inclusive.*

*(i) Draw the graph of probability density function.*

*(ii) Show that (a)* $P(B \mid A) = \frac{2}{3}$, *(b)* $P(\bar{A} \cap \bar{B}) = \frac{1}{3}$

**Solution.** *(i)* The graph of p.d.f. is given below.




*(ii) (a)* $P(A) = \int_0^2 f(x)\,dx = \int_0^1 \frac{1}{9}(x+1)\,dx + \int_1^{3/2} \frac{4}{9}\left(x - \frac{1}{2}\right)dx + \int_{3/2}^{2} \frac{4}{9}\left(\frac{5}{2} - x\right)dx$

$$= \frac{1}{2} \quad \text{( on simplification )}$$

$$P(A \cap B) = P(1 \le X \le 2) = \int_1^2 f(x)\,dx$$

$$= \int_1^{3/2} \frac{4}{9}\left(x - \frac{1}{2}\right)dx + \int_{3/2}^{2} \frac{4}{9}\left(\frac{5}{2} - x\right)dx$$

$$= \frac{4}{9}\left[\frac{x^2}{2} - \frac{x}{2}\right]_1^{3/2} + \frac{4}{9}\left[\frac{5}{2}x - \frac{x^2}{2}\right]_{3/2}^{2} = \frac{1}{3}$$

( on simplification )

$$\therefore \quad P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{1/3}{1/2} = \frac{2}{3}$$

*(b)* $\bar{A} \cap \bar{B}$ means that waiting time is more than 3 minutes.

$$\therefore \quad P(\bar{A} \cap \bar{B}) = P(X > 3) = \int_3^{\infty} f(x)\,dx = \int_3^6 f(x)\,dx + \int_6^{\infty} f(x)\,dx$$

$$= \int_3^6 \frac{1}{9}\,dx = \frac{1}{9} \left| x \right|_3^6 = \frac{1}{3}$$

**Example 5·13.** *The amount of bread (in hundreds of pounds) X that a certain bakery is able to sell in a day is found to be a numerical valued random phenomenon, with a probability function specified by the probability density function $f(x)$, given by*

$$f(x) = A\,.\,x, \qquad \text{for } 0 \le x < 5$$
$$= A(10-x), \quad \text{for } 5 \le x < 10$$
$$= 0, \quad \textit{otherwise}$$

*(a) Find the value of A such that $f(x)$ is a probability density function.*

*(b) What is the probability that the number of pounds of bread that will be sold tomorrow is*

*(i) more than 500 pounds,*

*(ii) less than 500 pounds,*

*(iii) between 250 and 750 pounds?* **[Agra Univ. B.Sc., 1989]**

*(c) Denoting by A, B, C the events that the pounds of bread sold are as in b (i), b (ii) and b (iii) respectively, find $P(A \mid B), P(A \mid C)$. Are (i) A and B independent events? (ii) Are A and C independent events ?*

**Solution.** *(a)* In order that $f(x)$ should be a probability density function

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

*i.e.,*
$$\int_0^5 A\,x\,dx + \int_5^{10} A(10-x)\,dx = 1$$

$$\Rightarrow \qquad A = \frac{1}{25} \qquad \text{(On simplification)}$$

*(b) (i)* The probability that the number of pounds of bread that will be sold tomorrow is more than 500 pounds, *i.e.,*

$$P(5 \le X \le 10) = \int_5^{10} \frac{1}{25}(10-x)\,dx = \frac{1}{25}\left| 10x - \frac{x^2}{2} \right|_5^{10}$$
$$= \frac{1}{25}\left(\frac{25}{2}\right) = \frac{1}{2} = 0{\cdot}5$$

*(ii)* The probability that the number of pounds of bread that will be sold tomorrow is less than 500 pounds, *i.e.,*

$$P(0 \le X \le 5) = \int_0^5 \frac{1}{25}\,.\,x\,dx = \frac{1}{25}\left|\frac{x^2}{2}\right|_0^5 = \frac{1}{2} = 0{\cdot}5$$

*(iii)* The required probability is given by

$$P(2{\cdot}5 \le X \le 7{\cdot}5) = \int_{2\cdot5}^{5} \frac{1}{25}\,x\,dx + \int_5^{7\cdot5} \frac{1}{25}(10-x)\,dx = \frac{3}{4}$$

*(c)* The events *A, B* and *C* are given by

$$A: 5 < X \le 10\,; \quad B: 0 \le X < 5\,; \quad C: 2{\cdot}5 < X < 7{\cdot}5$$

Then from parts *b (i), (ii)* and *(iii)*, we have

$$P(A) = 0{\cdot}5\ ,\ \ P(B) = 0{\cdot}5\ ,\ \ P(C) = \frac{3}{4}$$

The events $A \cap B$ and $A \cap C$ are given by

$$A \cap B = \phi \ \text{ and } \ A \cap C\ :\ 5 < X < 7{\cdot}5$$

$\therefore \qquad P(A \cap B) = P(\phi) = 0$

and $\qquad P(A \cap C) = \int_5^{7\cdot5} f(x)\,dx = \frac{1}{25}\int_5^{7\cdot5} (10 - x)\,dx$

$$= \frac{1}{25} \times \frac{75}{8} = \frac{3}{8}$$

$$P(A)\,.\,P(C) = \frac{1}{2} \times \frac{3}{4} = \frac{3}{8} = P(A \cap C)$$

$\Rightarrow \qquad A$ and $C$ are independent.

Again $\ P(A)\,.\,P(B) = \frac{1}{4} \neq P(A \cap B)$

$\Rightarrow \qquad A$ and $B$ are not independent.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = 0$$

$$P(A \mid C) = \frac{P(A \cap C)}{P(C)} = \frac{3/8}{3/4} = \frac{1}{2}$$

**Example 5·14.** *The mileage C in thousands of miles which car owners get with a certain kind of tyre is a random variable having probability density function*

$$f(x) = \frac{1}{20}\ e^{-x/20}\ ,\ \text{for } x > 0$$
$$= 0\ ,\ \text{for } x \leq 0$$

*Find the probabilities that one of these tyres will last*

*(i) at most 10,000 miles,*

*(ii) anywhere from 16,000 to 24,000 miles.*

*(iii) at least 30,000 miles.* **(Bombay Univ. B.Sc. 1989)**

**Solution.** Let *r.v.* $X$ denote the mileage (in '000 miles) with a certain kind of tyre. Then required probability is given by:

*(i)* $\qquad P(X \leq 10) = \int_0^{10} f(x)\,dx = \frac{1}{20}\int_0^{10} e^{-x/20}\,dx$

$$= \frac{1}{20}\left|\frac{e^{-x/20}}{-1/20}\right|_0^{10} = 1 - e^{-1/2}$$

$$= 1 - 0{\cdot}6065 = 0{\cdot}3935$$

*(ii)* $P(16 \le X \le 24) = \frac{1}{20} \int_{16}^{24} \exp\left(-\frac{x}{20}\right) dx = \left| -e^{-x/20} \right|_{16}^{24}$

$= e^{-16/20} - e^{-24/20} = e^{-4/5} - e^{-6/5}$

$= 0{\cdot}4493 - 0{\cdot}3012 = 0{\cdot}1481$

*(iii)* $P(X \ge 30) = \int_{30}^{\infty} f(x)\, dx = \frac{1}{20} \left| \frac{e^{-x/20}}{-1/20} \right|_{30}^{\infty}$

$= e^{-1.5} = 0{\cdot}2231$

## EXERCISE 5 (c)

1. *(a)* A continuous random variable $X$ follows the probability law

$$f(x) = Ax^2,\ 0 \le x \le 1$$

Determine $A$ and find the probability that *(i)* $X$ lies between 0·2 and 0·5, *(ii)* $X$ is less than 0·3, *(iii)* $1/4 < X < 1/2$ and *(iv)* $X > 3/4$ given $X > 1/2$.

**Ans.** $A = 0{\cdot}3$, *(i)* 0·117, *(ii)* 0·027, *(iii)* 15/256 and *(iv)* 27/56.

*(b)* If a random variable $X$ has the density function

$$f(x) = \begin{cases} 1/4, & -2 < x < 2 \\ 0, & elsewhere. \end{cases}$$

Obtain *(i)* $P(X < 1)$, *(ii)* $P(|X| > 1)$ *(iii)* $P(2X + 3 > 5)$

**(Kerala Univ. B.Sc., Sept.1992)**

**Hint.** *(ii)* $P(|X| > 1) = P(X > 1 \text{ or } X < -1) = \int_{-2}^{-1} f(x)\, dx + \int_{1}^{2} f(x)\, dx$

or $P(|X| > 1) = 1 - P(|X| \le 1) = 1 - P(-1 \le X \le 1)$

**Ans.** *(i)* 3/4, *(ii)* 1/2 *(iii)* 1/4.

2. Are any of the following probability mass or density functions? Prove your answer in each case.

*(a)* $f(x) = x;\ x = \frac{1}{16}, \frac{3}{16}, \frac{1}{4}, \frac{1}{2}$

*(b)* $f(x) = \lambda e^{-\lambda x};\ x \ge 0;\ \lambda > 0$

*(c)* $f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 4 - 2x, & 1 < x < 2 \\ 0, & elsewhere, \end{cases}$

**(Calicut Univ. B. Sc., Oct. 1989)**

**Ans.** *(a)* and *(b)* are p.m.f./p.d.f.'s, *(c)* is not.

3. If $f_1$ and $f_2$ are p.d.f.'s and $\theta_1 + \theta_2 = 1$, check if.

$$g(x) = \theta_1 f_1(x) + \theta_2 f_2(x), \quad \text{is a p.d.f.}$$

**Ans.** $g(x)$ is a p.d.f. if $0 \le (\theta_1, \theta_2) \le 1$ and $\theta_1 + \theta_2 = 1$.

4. A continuous random variable $X$ has the probability density function :

$$f(x) = A + Bx, \quad 0 \le x \le 1.$$

If the mean of the distribution is $\frac{1}{2}$, find $A$ and $B$.

**Hint :** Solve $\int_0^1 f(x)\, dx = 1$ and $\int_0^1 x\, f(x)\, dx = \frac{1}{2}$. Find $A$ and $B$.

5. For the following density function

$$f(x) = c\, x^2 (1 - x), \quad 0 < x < 1,$$

find *(i)* the constant $c$, and *(ii)* mean.

**[Calicut Univ. B.Sc.(subs.), 1991]**

**Ans.** *(i)* $c = 12$ ; *(ii)* mean = 3/5 .

6. A continuous distribution of a variable $X$ in the range $(-3, 3)$ is defined by

$$\begin{aligned} f(x) &= \tfrac{1}{16} (3 + x)^2, \quad -3 \le x \le -1 \\ &= \tfrac{1}{16} (6 - 2x^2), \quad -1 \le x \le 1 \\ &= \tfrac{1}{16} (3 - x)^2, \quad 1 \le x \le 3 \end{aligned}$$

*(i)* Verify that the area under the curve is unity.
*(ii)* Find the mean and variance of the above distribution.

**(Madras Univ. B.Sc., Oct. 1992; Gujarat Univ. B.Sc., Oct. 1986)**

**Hint:** $\int_{-3}^{3} f(x)\, dx = \int_{-3}^{-1} f(x)\, dx + \int_{-1}^{1} f(x)\, dx + \int_{1}^{3} f(x)\, dx$

**Ans.** Mean=0, Variance=1

7. If the random variable $X$ has the p.d.f.,

$$\begin{aligned} f(x) &= \tfrac{1}{2} (x + 1), \quad -1 < x < 1 \\ &= 0, \text{ elsewhere,} \end{aligned}$$

find the coefficient of skewness and kurtosis.

8. *(a)* A random variable $X$ has the probability density function given by

$$f(x) = 6x(1 - x), \quad 0 \le x \le 1$$

Find the mean $\mu$, mode and S.D. $\sigma$, Compute $P(\mu - 2\sigma < X < \mu + 2\sigma)$. Find also the mean deviation about the median.

**(Lucknow Univ. B.Sc., 1988)**

*(b)* For the continuous distribution

$$dF = y_o (x - x^2)\, dx \;;\; 0 \le x \le 1, \; y_o \text{ being a constant.}$$

Find *(i)* arithmetic mean, *(ii)* harmonic mean, *(iii)* Median, *(iv)* Mode and *(v)* $r$th moment about mean. Hence find $\beta_1$ and $\beta_2$ and show that the distribution is symmetrical. **(Delhi Univ. B.Sc., 1992 ; Karnatak Univ. B.Sc., 1991)**

**Ans.** Mean = Median = Mode = $\frac{1}{2}$

(c) Find the mean, mode and median for the distribution,

$$dF(x) = \sin x\, dx, \quad 0 \le x \le \pi/2$$

**Ans.** 1, $\pi/2$, $\pi/3$

9. If the function $f(x)$ is defined by

$$f(x) = c\, e^{-\alpha x}, \quad 0 \le x < \infty, \quad \alpha > 0$$

(i) Find the value of constant $c$.

(ii) Evaluate the first four moments about mean.

**[Gauhati Univ. B.Sc. 1987]**

**Ans.** (i) $c = \alpha$, (ii) 0, $1/\alpha^2$, $2/\alpha^3$, $9/\alpha^4$.

10. (a) Show that for the exponential distribution

$$dP = y_o \,.\, e^{-x/\sigma}\, dx, \quad 0 \le x < \infty, \quad \sigma > 0$$

the mean and S.D. are both equal to $\sigma$ and that the interquartile range is $\sigma \log_e 3$. Also find $\mu_r'$ and show that $\beta_1 = 4$, $\beta_2 = 9$.

**[Agra Univ. B.Sc., 1986 ; Madras Univ. B.Sc., 1987]**

(b) Define the harmonic mean (H.M.) of variable $X$ as the reciprocal of the expected value of $1/X$, show that the H.M. of variable which ranges from 0 to $\infty$ with probability density $\frac{1}{6} x^3 e^{-x}$ is 3.

11. (a) Find the mean, variance and the co-efficients $\beta_1$, $\beta_2$ of the distribution,

$$dF = k\, x^2\, e^{-x}\, dx, \quad 0 < x < \infty.$$

**Ans.** $k = 1/2$; 3, 3, 4/3 and 5.

(b) Calculate $\beta_1$ for the distribution,

$$dF = k\, x\, e^{-x}\, dx, \quad 0 < x < \infty$$

**Ans.** 2 **[Delhi Univ. B.Sc. (Hons. Subs.), 1988]**

12. A continuous random variable $X$ has a p.d.f. given by

$$f(x) = k\, x\, e^{-\lambda x}, \quad x \ge 0, \quad \lambda > 0$$
$$= 0, \text{ otherwise}$$

Determine the constant $k$, obtain the mean and variance of $X$.

**[Nagpur Univ. B.Sc. 1990]**

13. For the probability density function,

$$f(x) = \frac{2(b+x)}{b(a+b)}, \quad -b \le x < 0$$
$$= \frac{2(a-x)}{a(a+b)}, \quad 0 \le x \le a$$

Find mean, median and variance. **[Calcutta Univ. B.Sc. 1984]**

**Ans.** Mean $= (a-b)/3$, Variance $= (a^2 + b^2 + ab)/18$,

Median $= a - \sqrt{a(a+b)/2}$

*(ii)* Show that, if terms of order $(a-b)^2/a^2$ are neglected, then

$$\text{mean} - \text{median} = (\text{mean} - \text{mode})/4$$

**14.** A variable $X$ can assume values only between 0 and 5 and the equation of its frequency curve is

$$y = A \sin \tfrac{1}{5} \pi x, \quad 0 \le x \le 5$$

where $A$ is a constant such that the area under the curve is unity. Determine the value of $A$ and obtain the median and quartiles of the distribution.

Show also that the variance of the distribution is $50\left\{\frac{1}{8} - \frac{1}{\pi^2}\right\}$.

**Ans.** 1/10; 2·5, 4/3, 10/3

**15.** A continuous variable $X$ is distributed over the interval [0, 1] with p.d.f. $ax^2 + bx$, where $a, b$ are constants. If the arithmetic mean of $X$ is 0·5, find the values of $a$ and $b$.

**Ans.** $-6, 6$

**16.** A man leaves his house at the same time every morning and the time taken to journey to work has the following probability density function : less than 30 minutes, zero, between 30 minutes and 60 minutes, uniform with density $k$ ; between 60 minutes and 70 minutes, uniform with density $2k$ ; and more than 70 minutes, zero. What is the probability that on one particular day he arrives at work later than on the previous day but not more than 5 minutes later.

**17.** The density function of sheer strength of spot welds is given by

$$\begin{aligned} f(x) &= x/160{,}000 \quad \text{for } 0 \le x \le 400 \\ &= (800 - x)/160{,}000 \quad \text{for } 400 \le x \le 800 \end{aligned}$$

Find the number $a$ such that

Prob. $(X < a) = 0{\cdot}50$ and the number $b$ such that

Prob. $(X < b) = 0{\cdot}90$. Find the mean, median and variance of $X$. .

[Delhi Univ. B.E., 1987]

**18.** A batch of small calibre ammunition is accepted as satisfactory if none of a sample of five shot falls more than 2 feet from the centre of the target at a given range. If $X$, the distance from the centre of the target to a given impact point, actually has the density

$$f(x) = k \cdot 2x e^{-x^2}, \quad 0 < x < 3$$

where $k$ is a number which makes it probability density function, what is the value of $k$ and what is the probability that the batch will be accepted?

[Nagpur Univ. B.E., 1987]

**Hint.** $\int_0^3 f(x)\, dx = 1 \quad \Rightarrow \quad k = 1/(1 - e^{-9})$

Reqd. Prob. = P [ Each of a sample of 5 shots falls within a distance of 2 ft. from the centre ]

$$= [P(0<X<2)]^5 = \left[\int_0^2 f(x)\,dx\right]^5 = \left[\frac{1-e^{-4}}{1-e^{-9}}\right]^5$$

**19.** A random variable $X$ has the p.d.f. :

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & otherwise \end{cases}$$

Find *(i)* $P\left(X < \frac{1}{2}\right)$, *(ii)* $P\left(\frac{1}{4} < X < \frac{1}{2}\right)$, *(iii)* $P\left(X > \frac{3}{4} \,\middle|\, X > \frac{1}{2}\right)$, and *(iv)* $P\left(X < \frac{3}{4} \,\middle|\, X > \frac{1}{2}\right)$. **(Gorakhpur Univ. B.Sc., 1988)**

**Ans.** *(i)* 1/4 , *(ii)* 3/16 , *(iii)* $\dfrac{P(X > 3/4)}{P(X > 1/2)} = \dfrac{7/16}{3/4} = \dfrac{7}{12}$ ; *(iv)* $\dfrac{P(1/2 < X < 3/4)}{P(X > 1/2)}$

**5·4·3. Continuous Distribution Function.** If $X$ is a continuous random variable with the p.d.f. $f(x)$, then the function

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt, \quad -\infty < x < \infty. \qquad \text{...(5·12)}$$

is called the *distribution function* (d.f.) or sometimes the *cumulative distribution function* (c.d.f.) of the random variable $X$.

**Remarks 1.** $0 \le F(x) \le 1, \quad -\infty < x < \infty.$

**2.** From analysis (Riemann integral), we know that

$$F'(x) = \frac{d}{dx} F(x) = f(x) \ge 0 \qquad [\because f(x) \text{ is p.d.f.}]$$

$\Rightarrow$ $F(x)$ is non-decreasing function of $x$.

**3.** $$F(-\infty) = \lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{-\infty} f(x)\,dx = 0$$

and $$F(+\infty) = \lim_{x \to \infty} F(x) = \lim_{x \to \infty} \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{\infty} f(x)\,dx = 1$$

**4.** $F(x)$ is a continuous function of $x$ on the right.

**5.** The discontinuities of $F(x)$ are at the most countable.

**6.** It may be noted that

$$P(a \le X \le b) = \int_a^b f(x)\,dx = \int_{-\infty}^{b} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx$$

$$= P(X \le b) - P(X \le a) = F(b) - F(a)$$

Similarly

$$P(a < X < b) = P(a < X \le b) = P(a \le X < b) = \int_a^b f(t)\,dt$$

7. Since $F'(x) = f(x)$, we have

$$\frac{d}{dx} F(x) = f(x) \quad \Rightarrow \quad dF(x) = f(x)\,dx$$

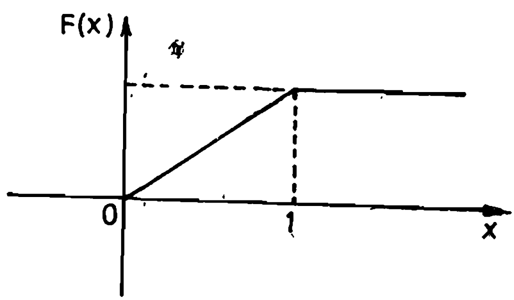
This is known as probability differential of $X$.

**Remarks.** 1. It may be pointed out that the properties (2), (3) and (4) above uniquely characterise the distribution functions. This means that any function $F(x)$ satisfying (2) to (4) is the distribution function of some random variable, and any function $F(x)$ violating any one or more of these three properties cannot be the distribution function of any random variable.

2. Often, one can obtain a p.d.f. from a distribution function $F(x)$ by differentiating $F(x)$, provided the derivative exists. For example, consider

$$F_X(x) = \begin{cases} 0, & \text{for } x < 0 \\ x, & \text{for } 0 \le x \le \\ 1, & \text{for } x > 1 \end{cases}$$

The graph of $F(x)$ is given by bold lines. Obviously we see that $F(x)$ is continuous from right as stipulated in (4) and we also see that $F(x)$ is not continuous at $x = 0$ and $x = 1$ and hence is not derivable at $x = 0$ and $x = 1$.




Differentiating $F(x)$ w.r.t. $x$, we get

$$\frac{d}{dx} F(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \textit{otherwise} \end{cases}$$

[Note the strict inequality in $0 < x < 1$, since $F(x)$ is not derivable at $x = 0$ and $x = 1$]

Let us define

$$f(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \textit{otherwise} \end{cases}$$

Then $f(x)$ is a p.d.f. for $F$.

**Example 5·15.** *Verify that the following is a distribution function:*

$$F(x) = \begin{cases} 0, & x < -a \\ \dfrac{1}{2}\left(\dfrac{x}{a} + 1\right), & -a \le x \le a \\ 1, & x > a \end{cases}$$

**(Madras Univ. B.Sc., 1992)**

**Solution.** Obviously the properties *(i), (ii), (iii)* and *(iv)* are satisfied. Also we observe that $F(x)$ is continuous at $x = a$ and $x = -a$, as well.

Now

$$\frac{d}{dx} F(x) = \begin{cases} \dfrac{1}{2a}, & -a \le x \le a \\ 0, & otherwise \end{cases}$$

$$= f(x), \text{ say}$$

In order that $F(x)$ is a distribution function, $f(x)$ must be a p.d.f. Thus we have to show that

$$\int_{-\infty}^{\infty} f(x)\, dx = 1$$

Now $\quad \displaystyle\int_{-\infty}^{\infty} f(x)\, dx = \int_{-a}^{a} f(x)\, dx = \frac{1}{2a}\int_{-a}^{a} 1\,.\, dx = 1$

Hence $F(x)$ is a d.f.

**Example 5·16.** *Suppose the life in hours of a certain kind of radio tube has the probability density function :*

$$f(x) = \frac{100}{x^2}, \quad when\ x \ge 100$$
$$= 0, \quad when\ x < 100$$

*Find the distribution function of the distribution. What is the probabilty that none of three such tubes in a given radio set will have to be replaced during the first 150 hours of operation ? What is the probability that all three of the original tubes will have been replaced during the first 150 hours ?* **(Delhi Univ. B.Sc, Oct. 1988)**

**Solution.** Probability that a tube will last for first 150 hours is given by

$$P(X \le 150) = P(0 < X < 100) + P(100 \le X \le 150)$$
$$= \int_{100}^{150} f(x)\, dx = \int_{100}^{150} \frac{100}{x^2}\, dx = \frac{1}{3}$$

Hence the probability that none of the three tubes will have to be replaced during the first 150 hours is $(1/3)^3 = 1/27$.

The probability that a tube will not last for the first 150 hours is $1 - \frac{1}{3} = \frac{2}{3}$.

Hence the probability that all three of the original tubes will have to be replaced during the first 150 hours is $(2/3)^3 = 8/27$.

**Example 5·17.** *Suppose that the time in minutes that a person has to wait at a certain station for a train is found to be a random phenomenon, a probability function specified by the distribution function,*

$$\begin{aligned} F(x) &= 0, && \text{for } x \le 0 \\ &= \frac{x}{2}, && \text{for } 0 \le x < 1 \\ &= \frac{1}{2}, && \text{for } 1 \le x < 2 \\ &= \frac{x}{4}, && \text{for } 2 \le x < 4 \\ &= 1, && \text{for } x \ge 4 \end{aligned}$$

*(a) Is the Distribution Function continuous ? If so, give the formula for its probability density function ?*

*(b) What is the probability that a person will have to wait (i) more than 3 minutes, (ii) less than 3 minutes, and (iii) between 1 and 3 minutes ?*

*(c) What is the conditional probability that the person will have to wait for a train for (i) more than 3 minutes, given that it is more than 1 minute, (ii) less than 3 minutes given that it is more than 1 minute ?* **(Calicut Univ. B.Sc., 1985)**

**Solution.** *(a)* Since the value of the distribution function is the same at the points $x = 0, x = 1, x = 2,$ and $x = 4$ given by the two forms of $F(x)$ for $x < 0$ and $0 \le x < 1$, $0 \le x < 1$ and $1 \le x < 2$, $1 \le x < 2$ and $2 \le x < 4$, $2 \le x < 4$ and $x \ge 4$, the distribution function is continuous.

Probability density function $= f(x) = \dfrac{d}{dx} F(x)$

$$\therefore \quad \begin{aligned} f(x) &= 0, && \text{for } x < 0 \\ &= \frac{1}{2}, && \text{for } 0 \le x < 1 \\ &= 0, && \text{for } 1 \le x < 2 \\ &= \frac{1}{4}, && \text{for } 2 \le x < 4 \\ &= 0, && \text{for } x \ge 4 \end{aligned}$$

*(b)* Let the random variable $X$ represent the waiting time in minutes. Then

*(i)* Required probability $= P(X > 3) = 1 - P(X \le 3) = 1 - F(3)$

$$= 1 - \frac{1}{4} \cdot 3 = \frac{1}{4}$$

*(ii)* Required probability $= P(X < 3) = P(X \le 3) - P(X = 3)$

$$= F(3) = \frac{3}{4}$$

(Since, the probability that a continuous variable takes a fixed value is zero)

(iii) Required Probability $= P(1 < X < 3) = P(1 < X \le 3)$

$$= F(3) - F(1) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$$

(c) Let $A$ denote the event that a person has to wait for more than 3 minutes and $B$ the event that he has to wait for more than 1 minute. Then

$$P(A) = P(X > 3) = \frac{1}{4} \qquad [\textit{c.f.} \text{ (b), (i)}]$$

$$P(B) = P(X > 1) = 1 - P(X \le 1) = 1 - F(1) = 1 - \frac{1}{2} = \frac{1}{2}$$

$$P(A \cap B) = P(X > 3 \cap X > 1) = P(X > 3) = \frac{1}{4}$$

(i) Required probability is

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{1/2} = \frac{1}{2}$$

(ii) Required probability $= P(\bar{A} \mid B) = \dfrac{P(\bar{A} \cap B)}{P(B)}$

Now $P(\bar{A} \cap B) = P(X \le 3 \cap X > 1) = P(1 < X \le 3) = F(3) - F(1) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$

$$\therefore \quad P(\bar{A} \mid B) = \frac{1/4}{1/2} = \frac{1}{2}$$

**Example 5·18.** *A petrol pump is supplied with petrol once a day. If its daily volume X of sales in thousands of litres is distributed by*

$$f(x) = 5(1 - x)^4, \quad 0 \le x \le 1,$$

*what must be the capacity of its tank in order that the probability that its supply will be exhausted in a given day shall be 0·01 ?* **(Madras Univ. B.E., 1986)**

**Solution.** Let the capacity of the tank ( in '000 of litres) be '$a$' such that

$$P(X \ge a) = 0{\cdot}01 \quad \Rightarrow \quad \int_a^1 f(x)\,dx = 0{\cdot}01$$

$$\Rightarrow \quad \int_a^1 5(1 - x)^4\,dx = 0{\cdot}01 \quad \text{or} \quad \left[5 \cdot \frac{(1-x)^5}{(-5)}\right]_a^1 = 0{\cdot}01$$

$$\Rightarrow \quad (1 - a)^5 = 1/100 \quad \text{or} \quad 1 - a = (1/100)^{1/5}$$

$$\therefore \quad a = 1 - (1/100)^{1/5} = 1 - 0{\cdot}3981 = 0{\cdot}6019$$

Hence the capacity of the tank $= 0{\cdot}6019 \times 1000$ *litres* $= 601{\cdot}9$ *litres.*

**Example 5·19.** *Prove that mean deviation is least when measured from the median.* **[Delhi Univ. B.Sc. (Maths. Hons.), 1989]**

**Solution.** If $f(x)$ is the probability function of a random variable $X$, $a \le X \le b$, then mean deviation $M(A)$, say, about the point $x = A$ is given by

$$M(A) = \int_a^b |x - A|\, f(x)\,dx$$

$$= \int_a^A |x - A| f(x)\, dx + \int_A^b |x - A| f(x)\, dx$$

$$= \int_a^A (A - x) f(x)\, dx + \int_A^b (x - A) f(x)\, dx \qquad ...(1)$$

We want to find the value of '$A$' so that $M(A)$ is minimum. From the principle of maximum and minimum in differential calculus, $M(A)$ will be minimum for variations in $A$ if

$$\frac{\partial M(A)}{\partial A} = 0 \quad \text{and} \quad \frac{\partial^2 M(A)}{\partial A^2} > 0 \qquad ...(2)$$

Differentiating (1) w.r.t. '$A$' under the integral sign, since the functions $(A - x) f(x)$ and $(x - A) f(x)$ vanish at the point $x = A^*$, we get

$$\frac{\partial M(A)}{\partial A} = \int_a^A f(x)\, dx - \int_A^b f(x)\, dx \qquad ...(3)$$

Also $\quad \dfrac{\partial M(A)}{\partial A} = \int_a^A f(x)\, dx - \left[1 - \int_a^A f(x)\, dx\right],$

$$\left[\because \int_a^b f(x)\, dx = 1\right]$$

$$= 2\int_a^A f(x)\, dx - 1 = 2F(A) - 1,$$

where $F(\cdot)$ is the distribution function of $X$. Differentiating again w.r.t. $A$, we get

$$\frac{\partial^2}{\partial A^2} M(A) = 2f(A) \qquad ...(4)$$

Now $\dfrac{\partial M(A)}{\partial A} = 0$, on using (3) gives

$$\int_a^A f(x)\, dx = \int_A^b f(x)\, dx$$

*i.e.*, $A$ is the median value.

Also from (4), we see that

$$\frac{\partial^2 M(A)}{\partial A^2} > 0,$$

assuming that $f(x)$ does not vanish at the median value. Thus mean deviation is least when taken from median.

---

*If $f(x, \theta)$ is a continuous function of both variables $x$ and $\theta$, possessing continuous partial derivatives $\dfrac{\partial^2 f}{\partial x \partial \theta}$, $\dfrac{\partial^2 f}{\partial \theta \partial x}$ and $a$ and $b$ are differentiable functions of $\theta$, then

$$\frac{\partial}{\partial \theta}\left[\int_a^b f(x, \theta)\, dx\right] = \int_a^b \frac{\partial f}{\partial \theta}\, dx + f(b, \theta)\frac{db}{d\theta} - f(a, \theta)\frac{da}{d\theta}$$

## EXERCISE 5 (d)

1. *(a)* Explain the terms *(i)* probability differential, *(ii)* probability density function, and *(iii)* distribution function.

*(b)* Explain what is meant by a random variable. Distinguish between a discrete and a continuous random variable. Define distribution function of a random variable and show that it is monotonic non-decreasing everywhere and continuous on the right at every point.

[Madras Univ. B.Sc. (Stat Main), 1987]

*(c)* Show that the distribution function $F(x)$ of a random variable $X$ is a non-decreasing function of $x$. Determine the jump of $F(x)$ at a point $x_0$ of its discontinuity in terms of the probability that the random variable has the value $x_0$.

[Calcutta Univ. B.Sc. (Hons.), 1984]

2. The length (in hours) $X$ of a certain type of light bulb may be supposed to be a continuous random variable with probability density function :

$$f(x) = \frac{a}{x^3}, \quad 1500 < x < 2500$$
$$= 0, \text{ elsewhere.}$$

Determine the constant $a$, the distribution function of $X$, and compute the probability of the event $1{,}700 \leq X \leq 1{,}900$.

**Ans.** $a = 70{,}31{,}250$ ; $F(x) = \dfrac{a}{2}\left(\dfrac{1}{22{,}50{,}000} - \dfrac{1}{x^2}\right)$ and

$$P(1{,}700 < X < 1{,}900) = F(1{,}900) - F(1{,}700) = \frac{a}{2}\left(\frac{1}{28{,}90{,}000} - \frac{1}{36{,}10{,}000}\right)$$

3. Define the "distribution function" (or cumulative distribution function) of a random variable and state its essential properties.

Show that, whatever the distribution function $F(x)$ of a random variable $X$, $P[a \leq F(x) \leq b] = b - a$, $0 \leq a, b \leq 1$.

4. *(a)* The distribution function of a random variable $X$ is given by

$$F(x) = \begin{cases} 1 - (1 + x)e^{-x}, & \text{for } x \geq 0 \\ 0, & \text{for } x < 0 \end{cases}$$

Find the corresponding density function of random variable $X$.

*(b)* Consider the distribution for $X$ defined by

$$F(x) = \begin{cases} 0, & \text{for } x < 0 \\ 1 - \frac{1}{4} e^{-x} & \text{for } x \geq 0 \end{cases}$$

Determine $P(x = 0)$ and $P(x > 0)$.

[Allahabad Univ. B.Sc., 1992]

5. *(a)* Let $X$ be a continuous random variable with probability density function given by

$$f(x) = \begin{cases} ax, & 0 \le x \le 1 \\ a, & 1 \le x \le 2 \\ -ax + 3a, & 2 \le x \le 3 \\ 0, & elsewhere \end{cases}$$

*(i)* Determine the constant $a$.

*(ii)* Determine $F(x)$, and sketch its graph.

*(iii)* If three independent observations are made, what is the probability that exactly one of these three numbers is larger than $1 \cdot 5$ ?

**[Rajasthan Univ. M.Sc., 1987]**

**Ans.** *(i)* 1/2, *(iii)* 3/8.

*(b)* For the density $f_X(x) = k\, e^{-ax}(1 - e^{-ax})\ I_{0,\infty}(x)$, find the normalising constant $k$, $f_X(x)$ and evaluate $P(X > 1)$.

**[Delhi Univ. B.Sc. (Maths Hons.), 1989]**

**Ans.** $k = 2a$ ; $F(x) = 1 - 2e^{-ax} + e^{-2ax}$ ; $P(X > 1) = 2e^{-a} - e^{-2a}$

6. A random variable $X$ has the density function :

$$f(x) = K \cdot \frac{1}{1 + x^2}, \quad \text{if } -\infty < x < \infty$$
$$= 0, \quad otherwise$$

Determine $K$ and the distribution function.

Evaluate the probability $P(X \ge 0)$. Find also the mean and variance of $X$.

**[Karnatak Univ. B.Sc. 1985]**

**Ans.** $K = 1$, $F(x) = \dfrac{1}{\pi}\left\{\tan^{-1} x + \dfrac{\pi}{2}\right\}$, $P(x \ge 0) = 1/2$, Mean $= 0$, Variance does not exist.

7. A continuous random variable $X$ has the distribution function

$$F(x) = \begin{cases} 0, & \text{if } x \le 1 \\ k(x-1)^4, & \text{if } 1 < x \le 3 \\ 1, & \text{if } x > 3 \end{cases}$$

Find *(i)* $k$, *(ii)* the probability density function $f(x)$, and *(iii)* the mean and the median of $X$.

**Ans.** *(i)* $k = \dfrac{1}{16}$, *(ii)* $f(x) = \dfrac{1}{4}(x-1)^3$, $1 \le x \le 3$

8. Given $f(x) = \begin{cases} kx(1-x), & \text{for } 0 < x < 1 \\ 0, & elsewhere \end{cases}$

Show that

*(i)* $k = 1/5$, *(ii)* $F(x) = 0$ for $x \le 0$ and $F(x) = 1 - e^{-x/5}$, for $x > 0$

Using $F(x)$, show that

*(iii)* $P(3 < X < 5) = 0{\cdot}1809$, *(iv)* $P(X < 4) = 0{\cdot}5507$, *(v)* $P(X > 6) = 0{\cdot}3012$

**9.** A bombing plane carrying three bombs flies directly above a railroad track. If a bomb falls within 40 feet of track, the track will be sufficiently damaged to disrupt the traffic. Within a certain bomb site the points of impact of a bomb have the probability density function :

$$\begin{aligned} f(x) &= (100 + x)/10{,}000, \quad \text{when } -100 \le x \le 0 \\ &= (100 - x)/10{,}000, \quad \text{when } 0 \le x \le 100 \\ &= 0, \quad \text{elsewhere} \end{aligned}$$

where $x$ represents the vertical deviation (in feet) from the aiming point, which is the track in this case. Find the distribution function. If all the bombs are used, what is the probability that track will be damaged ?

**Hint.** Probability that track will be damaged by the bomb is given by

$$\begin{aligned} P(|X| < 40) &= P(-40 < X < 40) \\ &= \int_{-40}^{0} f(x)\,dx + \int_{0}^{40} f(x)\,dx \\ &= \int_{-40}^{0} \frac{100 + x}{10{,}000}\,dx + \int_{0}^{40} \frac{100 - x}{10{,}000}\,dx = \frac{16}{25} \end{aligned}$$

$\therefore$ Probability that a bomb will not damage the track $= 1 - \frac{16}{25} = \frac{9}{25}$

Probability that none of the three bombs damages the track

$$= \left(\frac{9}{25}\right)^3 = 0{\cdot}046656$$

Required probability that the track will be damaged $= 1 - 0{\cdot}046656 = 0{\cdot}953344$.

**10.** The length of time (in minutes) that a certain lady speaks on the telephone is found to be random phenomenon, with a probability function specified by the probability density function $f(x)$ as

$$\begin{aligned} f(x) &= A\,e^{-x/5}, \quad \text{for } x \ge 0 \\ &= 0, \quad \text{otherwise} \end{aligned}$$

*(a)* Find the value of $A$ that makes $f(x)$ a p.d.f.

**Ans.** $A = 1/5$

*(b)* What is the probability that the number of minutes that she will talk over the phone is

*(i)* More than 10 minutes, *(ii)* less than 5 minutes, and *(iii)* between 5 and 10 minutes ? **[Shivaji Univ. B.Sc., 1990]**

**Ans.** *(i)* $\dfrac{1}{e^2}$, *(ii)* $\dfrac{e-1}{e}$, *(iii)* $\dfrac{e-1}{e^2}$.

11. The probability that a person will die in the time interval $(t_1, t_2)$ is given by

$$A \int_{t_1}^{t_2} f(t)\, dt ;$$

where $A$ is a constant and the function $f(t)$ determined from long records, is

$$f(t) = \begin{cases} t^2 (100 - t)^2, & 0 \le t \le 100 \\ 0, & \textit{elsewhere} \end{cases}$$

Find the probability that a person will die between the ages 60 and 70 assuming that his age is $\geq 50$. **[Calcutta Univ. B.A. (Hons.), 1987]**

**5·5. Joint Probability Law.** Two random variables $X$ and $Y$ are said to be jointly distributed if they are defined on the same probability space. The sample points consist of 2–tuples. If the joint probability function is denoted by $P_{XY}(x, y)$ then the probability of a certain event $E$ is given by

$$P_{XY}(x, y) = P[(X, Y) \in E] \qquad \text{...(5·13)}$$

$(X, Y)$ is said to belong to $E$, if in the 2 dimensional space the 2–tuples lie in the Borel set $B$, representing the event $E$.

**5·5·1. Joint Probability Mass Function and Marginal and Conditional Probability Functions.** Let $X$ and $Y$ be random variables on a sample space $S$ with respective image sets $X(S) = \{x_1, x_2, \ldots, x_n\}$ and $Y(S) = \{y_1, y_2, \ldots, y_m\}$. We make the product set

$$X(S) \times Y(S) = \{x_1, x_2, \ldots, x_n\} \times \{y_1, y_2, \ldots, y_m\}$$

into a probability space by defining the probability of the ordered pair $(x_i, y_j)$ to be $P(X = x_i, Y = y_j)$ which we write $p(x_i, y_j)$. The function $p$ on $X(S) \times Y(S)$ defined by

$$p_{ij} = P(X = x_i \cap Y = y_j) = p(x_i, y_j) \qquad \text{...(5·14)}$$

is called the *joint probability function* of $X$ and $Y$ and is usually represented in the form of the following table :

| $X \backslash Y$ | $y_1$ | $y_2$ | $y_3$ | ... | $y_j$ | ... | $y_m$ | Total |
|---|---|---|---|---|---|---|---|---|
| $x_1$ | $p_{11}$ | $p_{12}$ | $p_{13}$ | ... | $p_{1j}$ | ... | $p_{1m}$ | $p_{1.}$ |
| $x_2$ | $p_{21}$ | $p_{22}$ | $p_{23}$ | ... | $p_{2j}$ | ... | $p_{2m}$ | $p_{2.}$ |
| $x_3$ | $p_{31}$ | $p_{32}$ | $p_{33}$ | ... | $p_{3j}$ | ... | $p_{3m}$ | $p_{3.}$ |
| $\vdots$ | $\vdots$ | | | | | | | $\vdots$ |
| $x_i$ | $p_{i1}$ | $p_{i2}$ | $p_{i3}$ | ... | $p_{ij}$ | ... | $p_{im}$ | $p_{i.}$ |
| $\vdots$ | $\vdots$ | | | | | | | $\vdots$ |
| $x_n$ | $p_{n1}$ | $p_{n2}$ | $p_{n3}$ | ... | $p_{nj}$ | ... | $p_{nm}$ | $p_{n.}$ |
| *Total* | $p_{.1}$ | $p_{.2}$ | $p_{.3}$ | ... | $p_{.j}$ | ... | $p_{.m}$ | 1 |

$$\therefore \quad \sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) = 1$$

Suppose the joint distribution of two random variables $X$ and $Y$ is given, then the probability distribution of $X$ is determined as follows :

$$p_X(x_i) = P(X = x_i) = P[X = x_i \cap Y = y_1] + P[X = x_i \cap Y = y_2] + \dots$$
$$+ P[X = x_i \cap Y = y_j] + \dots + P[X = x_i \cap \quad y_m]$$
$$= p_{i1} + p_{i2} + \dots + p_{ij} + \dots + p_{im}$$
$$= \sum_{j=1}^{m} p_{ij} = \sum_{j=1}^{m} p(x_i, y_j) = p_i. \qquad \dots (5\cdot 14\,a)$$

and is known as *marginal probability function of X.*

$$\text{Also} \quad \sum_{i=1}^{n} p_{i.} = p_{1.} + p_{2.} + \dots + p_{n.} = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = 1$$

Similarly, we can prove that

$$p_Y(y_j) = P(Y = y_j) = \sum_{i=1}^{n} p_{ij} = \sum_{i=1}^{n} p(x_i, y_j) = p_{.j} \qquad \dots (5\cdot 14\,b)$$

which is the *marginal probability function of Y.*

Also

$$P[X = x_i \mid Y = y_j] = \frac{P[X = x_i \cap Y = y_j]}{P[Y = y_j]} = \frac{p(x_i, y_j)}{p(y_j)} = \frac{p_{ij}}{p_{.j}}$$

This is known as *conditional probability function of X given Y =* $y_j$

Similarly

$$P[Y = y_j \mid X = x_i] = \frac{p(x_i, y_j)}{p(x_i)} = \frac{p_{ij}}{p_{i.}} \qquad \dots (5\cdot 14\,c)$$

is the *conditional probability function of Y given X =* $x_i$

$$\text{Also} \quad \sum_{i=1}^{n} \frac{p_{ij}}{p_{.j}} = \frac{p_{1j} + p_{2j} + \dots + p_{ij} + \dots + p_{nj}}{p_{.j}} = \frac{p_{.j}}{p_{.j}} = 1$$

Similarly

$$\sum_{j=1}^{n} \frac{p_{ij}}{p_{i.}} = 1$$

Two random variables $X$ and $Y$ are said to be *independent* if

$$P(X = x_i, Y = y_j) = P(X = x_i) \, . \, P(Y = y_j), \qquad \dots (5\cdot 14\,d)$$

otherwise they are said to be dependent.

**5·5·2. Joint Probability Distribution Function.** Let $(X, Y)$ be a two-dimensional random variable then their joint distribution function is denoted by $F_{XY}(x, y)$ and it represents the probability that simultaneously the observation

$(X, Y)$ will have the property $(X \le x$ and $Y \le y)$, *i.e.*,

$$F_{XY}(x, y) = P(-\infty < X \le x, -\infty < Y \le y)$$

$$= \int_{-\infty}^{x} \left[ \int_{-\infty}^{y} f_{XY}(x, y)\, dx\, dy \right. \qquad \ldots (5\cdot15)$$

(For continuous variables)

where $\qquad f_{XY}(x, y) \ge 0$

And $\displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y)\ dx\, dy = 1$ *or* $\displaystyle\sum_x \sum_y f(x, y) = 1$

**Properties of Joint Distribution Function**

**1.** (*i*) For the real numbers $a_1, b_1, a_2$ and $b_2$

$$P(a_1 < X \le b_1, a_2 < Y \le b_2) = F_{XY}(b_1, b_2) + F_{XY}(a_1, a_2) - F_{XY}(a_1, b_2) - F_{XY}(b_1, a_2)$$

[For proof, See Example 5·29]

(*ii*) Let $a_1 < a_2$, $b_1 < b_2$. We have

$(X \le a_1, Y \le a_2) + (a_1 < X \le b_1, Y \le a_2) = (X \le b_1, Y \le a_2)$

and the events on the L.H.S. are mutually exclusive. Taking probabilities on both-sides, we get :

$$F(a_1, a_2) + P(a_1 < X \le b_1, Y \le a_2) = F(b_1, a_2)$$

$$\Rightarrow \quad F(b_1, a_2) - F(a_1, a_2) = P(a_1 < X \le b_1, Y \le a_2)$$

$$\therefore \quad F(b_1, a_2) \ge F(a_1, a_2) \qquad [\text{Since } P(a_1 < X \le b_1, Y \le a_2) \ge 0]$$

Similarly it follows that

$$F(a_1, b_2) - F(a_1, a_2) = P(X \le a_1, a_2 < Y \le b_2) \ge 0$$

$$\Rightarrow \quad F(a_1, b_2) \ge F(a_1, a_2),$$

which shows that $F(x, y)$ is monotonic non-decreasing function.

**2.** $\quad F(-\infty, y) = 0 = F(x, -\infty)$, $F(+\infty, +\infty) = 1$

**3.** If the density function $f(x, y)$ is continuous at $(x, y)$ then

$$\frac{\partial^2 F}{\partial x\, \partial y} = f(x, y)$$

**5·5·3. Marginal Distribution Functions.** From the knowledge of joint distribution function $F_{XY}(x, y)$, it is possible to obtain the individual distribution functions, $F_X(x)$ and $F_Y(y)$ which are termed as marginal distribution function of $X$ and $Y$ respectively with respect to the joint distribution function $F_{XY}(x, y)$.

$$F_X(x) = P(X \le x) = P(X \le x, Y < \infty) = \lim_{y \to \infty} F_{XY}(x, y).$$

$$= F_{XY}(x, \infty) \qquad \ldots (5\cdot16)$$

Similarly, $F_Y(y) = P(Y \le y) = P(X < \infty, Y \le y)$

$$= \lim_{x \to \infty} F_{XY}(x, y) = F_{XY}(\infty, y)$$

$F_X(x)$ is termed as the marginal distribution function of $X$ corresponding to the joint distribution function $F_{XY}(x, y)$ and similarly $F_Y(y)$ is called marginal distribution function of the random variable $Y$ corresponding to the joint distribution function $F_{XY}(x, y)$.

In the case of jointly discrete random variables, the marginal distribution functions are given as

$$F_X(x) = \sum_y P(X \leq x, Y = y),$$

$$F_Y(y) = \sum_x P(X = x, Y \leq y)$$

Similarly in the case of jointly continuous random variable, the marginal distribution functions are given as

$$F_X(x) = \int_{-\infty}^{x} \left\{ \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy \right\} dx$$

$$F_Y(y) = \int_{-\infty}^{y} \left\{ \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx \right\} dy$$

**5·5·4 Joint Density Function, Marginal Density Functions.** From the joint distribution function $F_{XY}(x, y)$ of two dimensional continuous random variable we get the joint probabilty density function by differentiation as follows:

$$f_{XY}(x, y) = \partial^2 F(x, y)/\partial x \partial y$$

$$= \lim_{\delta x \to 0, \delta y \to 0} \frac{P(x \leq X \leq x + \delta x, y \leq Y \leq y + \delta y)}{\delta x\ \delta y}$$

Or it may be expressed in the following way also:

"The probability that the point $(x, y)$ will lie in the infinitesimal rectangular region, of area $dx\, dy$ is given by

$$P\left\{ x - \tfrac{1}{2} dx \leq X \leq x + \tfrac{1}{2} dx, y - \tfrac{1}{2} dy \leq Y \leq y + \tfrac{1}{2} dy \right\} = dF_{XY}(x, y)$$

and is denoted by $f_{XY}(x, y)\, dx\, dy$, where the function $f_{XY}(x, y)$ is called the joint probability density function of $X$ and $Y$.

The marginal probability function of $Y$ and $X$ are given respectively

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx \qquad \text{(for continuous variables)}$$

$$= \sum_x p_{XY}(x, y) \qquad \text{(for discrete variables)}$$

...(5·17)

and $$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy \qquad \text{(for continuous variables)}$$

$$= \sum_{y} p_{XY}(x, y) \qquad \text{(for discrete variables)}$$

(5·17a)

The marginal density functions of $X$ and $Y$ can be obtained in the following manner also.

$$\left.\begin{aligned} f_X(x) &= \frac{dF_X(x)}{dx} = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy \\ \text{and} \quad f_Y(y) &= \frac{dF_Y(y)}{dy} = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx \end{aligned}\right\} \qquad \text{... (5·17 b)}$$

**Important Remark.** If we know the joint p.d.f. (p.m.f.) $f_{XY}(x, y)$ of two random variables $X$ and $Y$, we can obtain the individual distributions of $X$ and $Y$ in the form of their marginal p.d.f.'s (p.m.f's) $f_X(x)$ and $f_Y(y)$ by using (5·17) and (5·17a). However, the converse is not true *i.e., from the marginal distributions of two jointly distributed random variables, we cannot determine the joint distributions of these two random variables.*

To verify this, it will suffice to show that two different joint p.m.f's (p.d.f.'s) have the same marginal distribution for $X$ and the same marginal distribution for $Y$. We give below two joint discrete probability distributions which have the same marginal distributions.

JOINT DISTRIBUTIONS HAVING SAME MARGINALS

***Probability Distribution I***

| Y \ X | 0 | 1 | $f_Y(y)$ |
|---|---|---|---|
| 0 | 0·28 | 0·37 | 0·65 |
| 1 | 0·22 | 0·13 | 0·35 |
| $f_X(x)$ | 0·50 | 0·50 | 1·00 |

***Probability Distribution II***

| Y \ X | 0 | 1 | $f_Y(y)$ |
|---|---|---|---|
| 0 | 0·35 | 0·30 | 0·65 |
| 1 | 0·15 | 0·20 | 0·35 |
| $f_X(x)$ | 0·50 | 0·50 | 1·00 |

As an illustration for continuous random variables, let $(X, Y)$ be continuous r.v. with joint p.d.f.

$$f_{XY}(x, y) = x + y \; ; \quad 0 \le (x, y) \le 1 \qquad \text{...(5·17 c)}$$

The marginal p.d.f.'s of $X$ and $Y$ are given by :

$$f_X(x) = \int_0^1 f(x, y)\, dy = \int_0^1 (x + y)\, dy = \left| xy + \frac{y^2}{2} \right|_0^1$$

$$\left.\begin{array}{lll} \Rightarrow \quad f_X(x) = x + \frac{1}{2} \ ; & 0 \le x \le 1 \\ \text{Similarly} \quad f_Y(y) = \int_0^1 f(x, y)\, dx = y + \frac{1}{2} \ ; & 0 \le y \le 1 \end{array}\right\} \quad \text{... (5·17 } d)$$

Consider another continuous joint p.d.f.

$$g(x, y) = \left(x + \tfrac{1}{2}\right)\left(y + \tfrac{1}{2}\right) \ ; \quad 0 \le (x, y) \le 1 \qquad \text{... (5·17 } e)$$

Then marginal p.d.f.'s of $X$ and $Y$ are given by :

$$g_1(x) = \int_0^1 g(x, y)\, dy = \left(x + \tfrac{1}{2}\right) \int_0^1 \left(y + \tfrac{1}{2}\right) dy$$

$$= \left(x + \tfrac{1}{2}\right) \left| \frac{y^2}{2} + \frac{1}{2} y \right|_0^1$$

$$\left.\begin{array}{lll} \Rightarrow \quad g_1(x) = x + \frac{1}{2} \ ; & 0 \le x \le 1 \\ \text{Similarly} \quad g_2(y) = y + \frac{1}{2} \ ; & 0 \le y \le 1 \end{array}\right\} \quad \text{... (5·17 } f)$$

(5·17 $d$) and (5·17 $f$) imply that the two joint p.d.f.'s in (5·17 $c$) and (5·17 $e$) have the same marginal p.d.f.'s (5·17 $d$) or (5·17 $f$).

Another illustration of continuous r.v.'s is given in Remark to Bivariate Normal Distribution, § 10·10·2.

**5·5·5. The Conditional Distribution Function and Conditinal Probability Density Function.** For two diamensional random variable $(X, Y)$, the joint distribution function $F_{XY}(x, y)$, for any real numbers $x$ and $y$ is given by

$$F_{XY}(x, y) = P(X \le x, Y \le y)$$

Now let $A$ be the event $(Y \le y)$ such that the event $A$ is said to occur when $Y$ assumes values up to and inclusive of $y$.

Using conditional probabilities we may now write

$$F_{XY}(x, y) = \int_{-\infty}^{x} P[A \mid X = x]\ dF_X(x) \qquad \text{... (5·18)}$$

The *conditional distribution function* $F_{Y|X}(y \mid x)$ denotes the distribution function of $Y$ when $X$ has already assumed the particular value $x$. Hence

$$F_{Y|X}(y \mid x) = P[Y \le y \mid X = x] = P[A \mid X = x]$$

Using this expression, the joint distribution function $F_{XY}(x, y)$ may be expressed in terms of the conditional distribution function as follows :

$$F_{XY}(x, y) = \int_{-\infty}^{x} F_{Y|X}(y \mid x)\ dF_X(x) \qquad \text{... (5·18 } a)$$

Similarly

$$F_{XY}(x, y) = \int_{-\infty}^{y} F_{X|Y}(x \mid y)\ dF_Y(y) \qquad \text{... (5·18 } b)$$

The *conditional probability density function* of $Y$ given $X$ for two random variables $X$ and $Y$ which are jointly continuously distributed is defined as follows, for two real numbers $x$ and $y$ :

$$f_{Y|X}(y|x) = \frac{\partial}{\partial y} F_{Y|X}(y|x) \quad \text{...(5·19)}$$

**Remarks :** 1. $f_X(x) > 0$, then

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}$$

**Proof.** We have

$$F_{XY}(x,y) = \int_{-\infty}^{x} F_{Y|X}(y|x)\, dF_X(x)$$

$$= \int_{-\infty}^{x} F_{Y|X}(y|x)\, f_X(x)\, dx$$

Differentiating w.r.t. $x$, we get

$$\frac{\partial}{\partial x} F_{XY}(x,y) = F_{Y|X}(y|x)\, f_X(x)$$

Differentiating w.r.t. $y$, we get

$$\frac{\partial}{\partial y}\left[\frac{\partial}{\partial x} F_{XY}(x,y)\right] = f_{Y|X}(y|x)\, f_X(x)$$

$$\Rightarrow \quad f_{XY}(x,y) = f_{Y|X}(y|x)\, f_X(x)$$

$$\Rightarrow \quad f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}$$

2. If $f_Y(y) > 0$, then

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

3. In terms of the differentials, we have

$$P(x < X \le x + dx \mid y < Y \le y + dy)$$

$$= \frac{P(x < X \le x + dx,\ y < Y \le y + dy)}{P(y < Y \le y + dy)}$$

$$= \frac{f_{XY}(x,y)\ dx\ dy}{f_Y(y)\ dy} = f_{X|Y}(x|y)\ dx$$

Whence $f_{X|Y}(x|y)$ may be interpreted as the conditional density function of $X$ on the assumption $Y = y$.

**5·5·6. Stochastic Independence.** Let us consider two random variables $X$ and $Y$ (of discrete or continuous type) with joint p.d.f. $f_{XY}(x,y)$ and marginal p.d.f.'s $f_X(x)$ and $g_Y(y)$ respectively. Then by the compound probability theorem

$$f_{XY}(x,y) = f_X(x)\ g_Y(y|x)$$

where $g_Y(y \mid x)$ is the conditional p.d.f. of $Y$ for given value of $X = x$.

If we assume that $g(y \mid x)$ does not depend on $x$, then by the definition of marginal p.d.f.'s, we get for continuous r.v.'s

$$g(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

$$= \int_{-\infty}^{\infty} f_X(x)\, g(y \mid x)\, dx$$

$$= g(y \mid x) \int_{-\infty}^{\infty} f_X(x)\, dx$$

[since $g(y \mid x)$ does not depend on $x$]

$$= g(y \mid x) \qquad [\because f(.) \text{ is p.d.f. of } X]$$

Hence

$$g(y) = g(y \mid x)$$

and $\quad f_{XY}(x, y) = f_X(x)\, g_Y(y)$ ... (*)

provided $g(y \mid x)$ does not depend on $x$. This motivates the following definition of independent random variables.

**Independent Random variables.** *Two r.v.'s $X$ and $Y$ with joint p.d.f. $f_{XY}(x, y)$ and marginal p.d.f.'s $f_X(x)$ and $g_Y(y)$ respectively are said to be stochastically independent if and only if*

$$f_{X,Y}(x, y) = f_X(x)\, g_Y(y) \qquad \text{... (5·20)}$$

**Remarks. 1.** In terms of the distribution function, we have the following definition :

*Two jointly distributed random variables $X$ and $Y$ are stochastically independent if and only if their joint distribution function $F_{X,Y}(\,.\,,\,.\,)$ is the product of their marginal distribution functions $F_X(\cdot)$ and $G_Y(\cdot)$, i.e., if for real $(x, y)$*

$$F_{X,Y}(x, y) = F_X(x)\, G_Y(y) \qquad \text{... (5·20 a)}$$

**2.** The variables which are not stochastically independent are said to be stochastically dependent.

**Theorem 5·8.** *Two random variables $X$ and $Y$ with joint p.d.f. $f(x, y)$ are stochastically independent if and only if $f_{X,Y}(x, y)$ can be expressed as the product of a non-negative function of $x$ alone and a non-negative function of $y$ alone, i.e., if*

$$f_{X,Y}(x, y) = h_X(x)\, k_Y(y) \qquad \text{... (5·21)}$$

*where* $h(\cdot) \geq 0$ and $k(\cdot) \geq 0$.

**Proof.** If $X$ and $Y$ are independent then by definition, we have

$$f_{X,Y}(x, y) = f_X(x) \cdot g_Y(y)$$

where $f(x)$ and $g(y)$ are marginal p.d.f. of $X$ and $Y$ respectively. Thus condition (5·21) is satisfied.

Conversely if (5·21) holds, then we have to prove that $X$ and $Y$ are independent. For continuous random variables $X$ and $Y$, the marginal p.d.f.'s are given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_{-\infty}^{\infty} h(x)\, k(y)\, dy$$

$$= h(x) \int_{-\infty}^{\infty} k(y)\, dy = c_1\, h(x), \text{ say} \qquad \ldots(*)$$

and $$g_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_{-\infty}^{\infty} h(x)\, k(y)\, dx$$

$$= k(y) \int_{-\infty}^{\infty} h(x)\, dx = c_2\, k(y), \text{ say} \qquad \ldots(**)$$

where $c_1$ and $c_2$ are constants independent of $x$ and $y$. Moreover

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$$

$$\Rightarrow \qquad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x)\, k(y)\, dx\, dy = 1$$

$$\Rightarrow \qquad \left(\int_{-\infty}^{\infty} h(x)\, dx\right)\left(\int_{-\infty}^{\infty} k(y)\, dy\right) = 1$$

$$\Rightarrow \qquad c_2\, c_1 = 1 \qquad \ldots(***)$$

Finally, we get

$$f_{X,Y}(x, y) = h_X(x)\, k_Y(y) = c_1 c_2\, h_X(x)\, k_Y(y) \qquad [\text{using } (***)]$$
$$= (c_1 h_X(x))\, (c_2 k_Y(y))$$
$$= f_X(x)\, g_Y(y) \qquad [\text{from } (*) \text{ and } (**)]$$

$\Rightarrow$ $X$ and $Y$ are stochastically independent.

**Theorem 5·9.** *If the random variables $X$ and $Y$ are stochastically independent, then for all possible selections of the corresponding pairs of real numbers $(a_1, b_1)$, $(a_2, b_2)$ where $a_i \le b_i$ for all $i = 1, 2$ and where the values $\pm\infty$ are allowed, the events $(a_1 < X \le b_1)$ and $(a_2 < Y \le b_2)$ are independent, i.e.,*

$$P[(a_1 < X \le b_1) \cap (a_2 < Y \le b_2)] = P(a_1 < X \le b_1)\, P(a_2 < Y \le b_2)$$

**Proof.** Since $X$ and $Y$ are stochastically independent, we have in the usual notations

$$f_{X,Y}(x, y) = f_X(x)\, g_Y(y) \qquad \ldots(*)$$

In case of continuous r.v.'s, we have

$$P[(a_1 < X \le b_1) \cap (a_2 < Y \le b_2)] = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y)\, dx\, dy$$

$$= \left( \int_{a_1}^{b_1} f_X(x)\ dx \right) \left( \int_{a_2}^{b_2} g_Y(y)\ dy \right) \qquad [\text{from } (*)]$$

$$= P(a_1 < X \le b_1)\, P(a_2 < Y \le b_2)$$

as desired.

**Remark.** In case of discrete r.v.'s theorems 5·8 and 5·9 can be proved on replacing integration by summation over the given range of the variables.

**Example 5.20.** *For the following bivariate probability distribution of X and Y, find*

*(i)* $P(X \le 1, Y = 2)$, *(ii)* $P(X \le 1)$, *(iii)* $P(Y = 3)$, *(iv)* $P(Y \le 3)$ *and (v)* $P(X < 3, Y \le 4)$

| X \ Y | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 0 | 0 | 0 | $\frac{1}{32}$ | $\frac{2}{32}$ | $\frac{2}{32}$ | $\frac{3}{32}$ |
| 1 | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ |
| 2 | $\frac{1}{32}$ | $\frac{1}{32}$ | $\frac{1}{64}$ | $\frac{1}{64}$ | 0 | $\frac{2}{64}$ |

**Solution.** The marginal distributions are given below :

| X \ Y | 1 | 2 | 3 | 4 | 5 | 6 | $p_X(x)$ |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | $\frac{1}{32}$ | $\frac{2}{32}$ | $\frac{2}{32}$ | $\frac{3}{32}$ | $\frac{8}{32}$ |
| 1 | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{10}{16}$ |
| 2 | $\frac{1}{32}$ | $\frac{1}{32}$ | $\frac{1}{64}$ | $\frac{1}{64}$ | 0 | $\frac{2}{64}$ | $\frac{8}{64}$ |
| $p_Y(y)$ | $\frac{3}{32}$ | $\frac{3}{32}$ | $\frac{11}{64}$ | $\frac{13}{64}$ | $\frac{6}{32}$ | $\frac{16}{64}$ | $\Sigma p(x) = 1$<br>$\Sigma p(y) = 1$ |

*(i)* $P(X \le 1, Y = 2) = P(X = 0, Y = 2) + P(X = 1, Y = 2)$

$$= 0 + \frac{1}{16} = \frac{1}{16}$$

*(ii)* $P(X \le 1) = P(X = 0) + P(X = 1)$

$$= \frac{8}{32} + \frac{10}{16} = \frac{7}{8} \qquad (\text{From above table})$$

*(iii)* $P(Y = 3) = \frac{11}{64}$ (From above table)

*(iv)* $P(Y \le 3) = P(Y = 1) + P(Y = 2) + P(Y = 3)$

$$= \frac{3}{32} + \frac{3}{32} + \frac{11}{64} = \frac{23}{64}$$

(v) $P(X<3, Y\le 4) = P(X=0, Y\le 4) + P(X=1, Y\le 4)$

$$+ P(X=2, Y\le 4)$$

$$= \left(\frac{1}{32}+\frac{2}{32}\right)+\left(\frac{1}{16}+\frac{1}{16}+\frac{1}{8}+\frac{1}{8}\right)$$

$$+\left(\frac{1}{32}+\frac{1}{32}+\frac{1}{64}+\frac{1}{64}\right)=\frac{9}{16}$$

**Example 5·21.** *The joint probability distribution of two random variables X and Y is given by :*

$$p(x, y) = \frac{2}{n(n+1)}, \quad \begin{array}{l} x = 1, 2, ..., n \\ y = 1, 2, ..., x \end{array}$$

*Examine whether X and Y are independent.* **(Calicut Univ. B.Sc., 1991)**

**Solution.** The joint probability distribution table along with the marginal distributions of $X$ and $Y$ is given below.

| $Y \backslash X$ | 1 | 2 | 3 | ......... | $n$ | $p_Y(y)$ |
|---|---|---|---|---|---|---|
| 1 | $\frac{2}{n(n+1)}$ | $\frac{2}{n(n+1)}$ | $\frac{2}{n(n+1)}$ | ......... | $\frac{2}{n(n+1)}$ | $\frac{2n}{n(n+1)}$ |
| 2 | – | $\frac{2}{n(n+1)}$ | $\frac{2}{n(n+1)}$ | ......... | $\frac{2}{n(n+1)}$ | $\frac{2(n-1)}{n(n+1)}$ |
| 3 | – | – | $\frac{2}{n(n+1)}$ | ......... | $\frac{2}{n(n+1)}$ | $\frac{2(n-2)}{n(n+1)}$ |
| ⋮ | ⋮ | | | | | |
| $n-1$ | – | – | – | $\frac{2}{n(n+1)}$ | $\frac{2}{n(n+1)}$ | $\frac{2\times 2}{n(n+1)}$ |
| n | – | – | – | – | $\frac{2}{n(n+1)}$ | $\frac{2}{n(n+1)}$ |
| $p_X(x)$ | $\frac{2}{n(n+1)}$ | $\frac{2\times 2}{n(n+1)}$ | $\frac{2\times 3}{n(n+1)}$ | ......... | $\frac{2\times n}{n(n+1)}$ | |

Note that $y = 1, 2, ..., x$.

When $x=1$, $y=1$; when $x=2$, $y=1,2$; when $x=3$, $y=1,2,3$ and so on.

From the above table, we see that

$$p_{XY}(x, y) \ne p_X(x)\, p_Y(y) \; ; \quad \forall \; x, y$$

$\Rightarrow$ $X$ and $Y$ are not independent.

**Example 5·22.** *Given the following bivariate probability distribution, obtain (i) marginal distributions of X and Y, (ii) the conditional distribution of X given Y = 2.*

| Y \ X | −1 | 0 | 1 |
|---|---|---|---|
| 0 | 1/15 | 2/15 | 1/15 |
| 1 | 3/15 | 2/15 | 1/15 |
| 2 | 2/15 | 1/15 | 2/15 |

(Mysore Univ. B.Sc., Oct. 1987)

**Solution.**

| Y \ X | −1 | 0 | 1 | $\sum_x p(x,y)$ |
|---|---|---|---|---|
| 0 | 1/15 | 2/15 | 1/15 | 4/15 |
| 1 | 3/15 | 2/15 | 1/15 | 6/15 |
| 2 | 2/15 | 1/15 | 2/15 | 5/15 |
| $\sum_y p(x,y)$ | 6/15 | 5/15 | 4/15 | 1 |

*(i) Marginal distribution of X.* From the above table, we get

$$P(X=-1)=\frac{6}{15}=\frac{2}{5}\ ;\ P(X=0)=\frac{5}{15}=\frac{1}{3}\ ;\ P(X=1)=\frac{4}{15}$$

*Marginal distribution of Y :*

$$P(Y=0)=\frac{4}{15}\ ;\ P(Y=1)=\frac{6}{15}=\frac{2}{5}\ ;\ P(Y=2)=\frac{5}{15}=\frac{1}{3}$$

*(ii) Conditional distribution of X given Y = 2.* We have

$$P(X=x\cap Y=2)=P(Y=2).P(X=x\mid Y=2)$$

$$\Rightarrow\qquad P(X=x\mid Y=2)=\frac{P(X=x\cap Y=2)}{P(Y=2)}$$

$$\therefore\qquad P(X=-1\mid Y=2)=\frac{P(X=-1\cap Y=2)}{P(Y=2)}=\frac{2/15}{1/3}=\frac{2}{5}$$

**Example 5·23.** *X and Y are two random variables having the joint density function, $f(x,y)=\frac{1}{27}(2x+y)$, where x and y can assume only the integer values 0, 1 and 2. Find the conditional distribution of Y for X = x.*

[South Gujarat Univ. B.Sc., 1988]

**Solution.** The joint probability function

$$f(x,y)=\frac{1}{27}(2x+y);\quad x=0,1,2;\quad y=0,1,2$$

gives the following table of joint probability distribution of $X$ and $Y$.

JOINT PROBABILITY DISTRIBUTION $f(x,y)$ OF $X$ AND $Y$

| $X\downarrow$ \ $Y\rightarrow$ | 0 | 1 | 2 | $f_X(x)$ |
|---|---|---|---|---|
| 0 | 0 | 1/27 | 2/27 | 3/27 |
| 1 | 2/27 | 3/27 | 4/27 | 9/27 |
| 2 | 4/27 | 5/27 | 6/27 | 15/27 |

For example $f(0,0)=\frac{1}{27}(0+2\times 0)=0$

$f(1,0)=\frac{1}{27}(0+2\times 1)=\frac{2}{27}$ ; $f(2,0)=\frac{1}{27}(0+2\times 2)=\frac{4}{27}$
and so on.

The marginal probability distribution of $X$ is given by

$$f_X(x)=\sum_y f(x,y),$$

and is tabulated in last column of above table.

The conditional distribution of $Y$ for $X=x$ is given by

$$f_{Y|X}(Y=y\mid X=x)=\frac{f(x,y)}{f_X(x)}$$

and is obtained in the following table.

CONDITIONAL DISTRIBUTION OF $Y$ FOR $X=x$

| $X$ \ $Y$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0 | 1/3 | 2/3 |
| 1 | 2/9 | 3/9 | 4/9 |
| 2 | 4/15 | 5/15 | 6/15 |

**Example 5·24.** *Two discrete random variables X and Y have the joint probability density function :*

$$p(x,y)=\frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!\,(x-y)!},\quad y=0,1,2,\ldots,x\ ;\ x=0,1,2,\ldots$$

*where $\lambda, p$ are constants with $\lambda > 0$ and $0 < p < 1$.*

*Find (i) The marginal probability density functions of X and Y.*

*(ii) The conditional distribution of Y for a given X and of X for a given Y.*

**(Poona Univ. B.Sc., 1986 ; Nagpur Univ. M.Sc., 1989)**

**Solution.** *(i)*

$$p_X(x) = \sum_{y=0}^{x} p(x,y) = \sum_{y=0}^{x} \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!}$$

$$= \frac{\lambda^x e^{-\lambda}}{x!} \sum_{y=0}^{x} \frac{x!\,p^y (1-p)^{x-y}}{y!(x-y)!} = \frac{\lambda^x e^{-\lambda}}{x!} \sum_{y=0}^{x} {}^xC_y p^y (1-p)^{x-y}$$

$$= \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots$$

which is the probability function of a Poisson distribution with parameter $\lambda$.

$$p_Y(y) = \sum_{x=0}^{\infty} p(x,y) = \sum_{x=y}^{\infty} \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!}$$

$$= \frac{(\lambda p)^y e^{-\lambda}}{y!} \sum_{x=y}^{\infty} \frac{[\lambda(1-p)]^{x-y}}{(x-y)!} = \frac{(\lambda p)^y e^{-\lambda}}{y!} e^{\lambda(1-p)}$$

$$= \frac{e^{-\lambda p}(\lambda p)^y}{y!}, \quad y = 0, 1, 2, \ldots,$$

which is the probability function of a Poisson distribution with parameter $\lambda p$.

*(ii)* The conditional distribution of $Y$ for given $X$ is

$$p_{Y|X}(y \mid x) = \frac{p_{XY}(x,y)}{p_X(x)} = \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y} x!}{y!(x-y)!\,\lambda^x e^{-\lambda}}$$

$$= \frac{x!}{y!(x-y)!} p^y (1-p)^{x-y} = {}^xC_y\, p^y (1-p)^{x-y}, \quad x > y$$

The conditional probability distribution of $X$ for given $Y$ is

$$p_{X|Y}(x \mid y) = \frac{p_{XY}(x,y)}{p_Y(y)}$$

$$= \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!} \cdot \frac{y!}{e^{-\lambda p}(\lambda p)^y} \qquad [\,c.f. \text{ Part (i)}\,]$$

$$= \frac{e^{-\lambda q}(\lambda q)^{x-y}}{(x-y)!}\,; \quad q = 1-p,\ x > y$$

**Example 5·25.** *The joint p.d.f. of two random variables X and Y is given by :*

$$f(x,y) = \frac{9(1+x+y)}{2(1+x)^4(1+y)^4}\,; \quad \begin{pmatrix} 0 \le x < \infty \\ 0 < y < \infty \end{pmatrix}$$

*Find the marginal distributions of X and Y, and the conditional distribution of Y for X = x.*

**Solution.** Marginal p.d.f. of $X$ is given by

$$f_X(x) = \int_0^\infty f(x, y)\, dy$$

$$= \frac{9}{2(1+x)^4} \int_0^\infty \frac{(1+y)+x}{(1+y)^4}\, dy$$

$$= \frac{9}{2(1+x)^4} \cdot \int_0^\infty \left[ (1+y)^{-3} + x\,(1+y)^{-4} \right] dy$$

$$= \frac{9}{2(1+x)^4} \left[ \left| \frac{-1}{2(1+y)^2} \right|_0^\infty + x \left| \frac{-1}{3(1+y)^3} \right|_0^\infty \right]$$

$$= \frac{9}{2(1+x)^4} \cdot \left[ \frac{1}{2} + \frac{x}{3} \right]$$

$$= \frac{3}{4} \cdot \frac{3+2x}{(1+x)^4} \; ; \; 0 < x < \infty$$

Since $f(x, y)$ is symmetric in $x$ and $y$, the marginal p.d.f. of $Y$ is given by

$$f_Y(y) = \int_0^\infty f(x, y)\, dx$$

$$= \frac{3}{4} \cdot \frac{3+2y}{(1+y)^4} \; ; \; 0 < y < \infty$$

The conditional distribution of $Y$ for $X = x$ is given by

$$f_{XY}(Y = y \mid X = x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

$$= \frac{9(1+x+y)}{2(1+x)^4 (1+y)^4} \cdot \frac{4(1+x)^4}{3(3+2x)}$$

$$= \frac{6(1+x+y)}{(1+y)^4 (3+2x)} \; ; \; 0 < y < \infty$$

**Example 5·26.** *The joint probability density function of a two-dimensional random variable (X,Y) is given by*

$$f(x, y) = 2 \; ; \; 0 < x < 1,\; 0 < y < x$$
$$= 0, \text{ elsewhere}$$

*(i) Find the marginal density functions of X and Y,*

*(ii) find the conditional density function of Y given X = x and conditional density function of X given Y = y, and*

*(iii) check for independence of X and Y.*

**[M.S.Baroda Univ. B.Sc., 1987; Karnataka Univ. B.Sc., Oct. 1988]**

**Solution.** Evidently $f(x, y) \geq 0$ and

$$\int_0^1 \int_0^x 2\, dx\, dy = 2 \int_0^1 x\, dx = 1$$

*(i)* The marginal p.d.f.'s of $X$ and $Y$ are given by

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy = \int_0^x 2\, dy = 2x, \quad 0 < x < 1$$

$$= 0, \text{ elsewhere}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx = \int_y^1 2dx = 2(1 - y), \quad 0 < y < 1$$

$$= 0, \text{ elsewhere}$$

*(ii)* The conditional density function of $Y$ given $X$ is

$$f_{Y|X}(y \mid x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{2}{2x} = \frac{1}{x}, \quad 0 < x < 1$$

The conditional density function of $X$ given $Y$ is

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{2}{2(1-y)} = \frac{1}{(1-y)}, \quad 0 < y < 1$$

*(iii)* Since $f_X(x)\, f_Y(y) = 2(2x)(1 - y) \neq f_{XY}(x, y)$, $X$ and $Y$ are not independent.

**Example 5·27.** *A gun is aimed at a certain point (origin of the coordinate system). Because of the random factors, the actual hit point can be any point (X,Y) in a circle of radius R about the origin. Assume that the joint density of X and Y is constant in this circle given by :*

$$f_{XY}(x, y) = k, \text{ for } x^2 + y^2 \leq R^2$$

$$= 0, \textit{ otherwise}$$

*(i) Compute k, (ii) show that*

$$f_X(x) = \frac{2}{\pi R} \left\{ 1 - \left( \frac{x}{R} \right)^2 \right\}^{1/2}, \quad \text{for } -R \leq x \leq R$$

$$= 0, \textit{ otherwise}$$

**[Calcutta Univ. B.Sc.(Stat. Hons.),1987]**

**Solution.** *(i)* The constant $k$ is computed from the consideration that the total probability is 1, *i.e.*,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1 \quad \Rightarrow \quad \iint_{x^2 + y^2 \leq R^2} k\, dx\, dy = 1$$

$$\Rightarrow \quad 4 \iint_I k\, dx\, dy = 1$$

where region $I$ is the first quadrant of the circle
$x^2 + y^2 = R^2$.

$$\Rightarrow \quad 4k \int_0^R \left( \int_0^{\sqrt{R^2 - x^2}} 1 \,.\, dy \right) dx = 1$$

$$\Rightarrow \quad 4k \int_0^R \sqrt{R^2 - x^2}\, dx = 1$$

$$\Rightarrow \quad 4k \left| x \sqrt{R^2 - x^2} + \frac{R^2}{2} \sin^{-1} \left( \frac{x}{R} \right) \right|_0^R = 1$$

$$\Rightarrow \quad 4k \cdot \left( \frac{R^2}{2} \cdot \frac{\pi}{2} \right) = 1 \quad \Rightarrow \quad k = \frac{1}{\pi R^2}$$

$$\therefore \quad f_{XY}(x, y) = 1/(\pi R^2) \; ; \quad x^2 + y^2 \le R^2$$
$$= 0, \quad \textit{otherwise}$$

*(ii)* $$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \frac{1}{\pi R^2} \int_{-\sqrt{R^2 - x^2}}^{\sqrt{R^2 - x^2}} 1 \cdot dy$$

$$\left[ \textit{because } x^2 + y^2 \le R^2 \Rightarrow -\left( R^2 - x^2 \right)^{1/2} \le y \le \left( R^2 - x^2 \right)^{1/2} \right]$$

$$= \frac{2}{\pi R^2} \int_0^{\sqrt{R^2 - x^2}} 1 \cdot dy = \frac{2}{\pi R^2} \left( R^2 - x^2 \right)^{1/2}$$

$$= \frac{2}{\pi R} \left[ 1 - \left( \frac{x}{R} \right)^2 \right]^{1/2}$$

**Example 5·28.** *Given:*

$$f(x, y) = e^{-(x+y)} \; I_{(0, \infty)}(x) \,.\, I_{(0, \infty)}(y),$$

*find* (i) $P(X > 1)$, (ii) $P(X < Y \mid X < 2Y)$, (iii) $P(1 < X + Y < 2)$

**[Delhi Univ. B.Sc. (Maths Hons.), 1987]**

**Solution.** We are given :

$$f(x, y) = e^{-(x+y)} \; ; \quad 0 \le x < \infty, \; 0 \le y < \infty \qquad \ldots (1)$$
$$= \left( e^{-x} \right) \left( e^{-y} \right)$$
$$= f_X(x) \,.\, f_Y(y) \; ; \quad 0 \le x < \infty, \; 0 \le y < \infty$$

$\Rightarrow$ $X$ and $Y$ are independent and

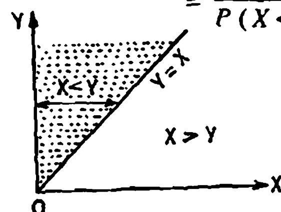
$$f_X(x) = e^{-x} \; ; \; x \ge 0 \quad \text{and} \quad f_Y(y) = e^{-y} \; ; \; y \ge 0 \qquad \ldots (2)$$

(i) $$P(X>1)=\int_1^{\infty} f_X(x)\,dx=\int_1^{\infty} e^{-x}\,dx$$

$$=\left|\frac{e^{-x}}{-1}\right|_1^{\infty}=\frac{1}{e}$$

(ii) $$P(X<Y \mid X<2Y)=\frac{P(X<Y \cap X<2Y)}{P(X<2Y)}$$

$$=\frac{P(X<Y)}{P(X<2Y)} \qquad \text{... (3)}$$




$$P(X<Y)=\int_0^{\infty}\left[\int_0^{y} f(x,y)\,dx\right]dy$$

$$=\int_0^{\infty}\left[e^{-y}\left|\frac{e^{-x}}{-1}\right|_0^{y}\right]dy=-\int_0^{\infty}e^{-y}\left(e^{-y}-1\right)dy$$

$$=-\left|\frac{e^{-2y}}{-2}+e^{-y}\right|_0^{\infty}=1-\frac{1}{2}=\frac{1}{2}$$

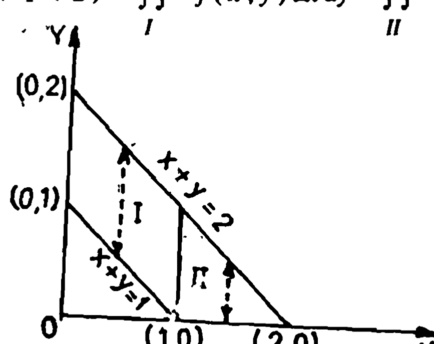
$$P(X<2Y)=\int_0^{\infty}\left[\int_0^{2y} f(x,y)\,dx\right]dy=-\int_0^{\infty}e^{-y}(e^{-2y}-1)\,dy$$

$$=-\left|\frac{e^{-3y}}{-3}+e^{-y}\right|_0^{\infty}=1-\frac{1}{3}=\frac{2}{3}$$

Substituting in (3),

$$P(X<Y \mid X<2Y)=\frac{1/2}{2/3}=\frac{3}{4}$$

(iii) $$P(1<X+Y<2)=\iint_{I} f(x,y)\,dx\,dy=\iint_{II} f(x,y)\,dx\,dy$$



$$= \int_0^1 \left( \int_{1-x}^{2-x} f(x,y)\,dy \right) dx + \int_1^2 \left( \int_0^{2-x} f(x,y)\,dy \right) dx$$

$$= \int_0^1 \left( e^{-x} \int_{1-x}^{2-x} e^{-y}\,dy \right) dx + \int_1^2 \left( e^{-x} \int_0^{2-x} e^{-y}\,dy \right) dx$$

$$= \int_0^1 \frac{e^{-x}}{-1}\left(e^{x-2} - e^{x-1}\right) dx + \int_1^2 \frac{e^{-x}}{-1}\left(e^{x-2} - 1\right) dx$$

$$= -\left(e^{-2} - e^{-1}\right) \int_0^1 1 \cdot dx - \int_1^2 \left(e^{-2} - e^{-x}\right) dx$$

$$= -\left(e^{-2} - e^{-1}\right) \Big| x \Big|_0^1 - \Big| e^{-2} \cdot x + e^{-x} \Big|_1^2$$

$$= 2/e - 3/e^2$$

**Example 5·29.** *(i) Let $F(x, y)$ be the d.f. of $X$ and $Y$. Show that*
$$P(a < X \le b,\ c < Y \le d) = F(b,d) - F(b,c) - F(a,d) + F(a,c)$$
*where $a, b, c, d$ are real constants $a < b$ ; $c < d$.*

*Deduce that if :* $F(x,y) = 1$, for $x + 2y \ge 1$
$F(x,y) = 0$, for $x + 2y < 1$,

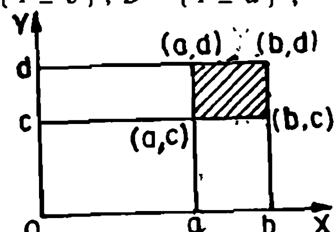
*then $F(x, y)$ cannot be joint distribution function of variables $X$ and $Y$.*

*(ii) Show that, with usual notation : for all $x$, $y$,*
$$F_X(x) + F_Y(y) - 1 \le F_{XY}(x,y) \le \sqrt{F_X(x) F_Y(y)}$$

**[Delhi Univ. B.Sc. (Maths Hons.), 1985 ]**

**Solution.** *(i)* Let us define the events :
$$A : \{X \le a\} \ ; \ B : \{X \le b\} \ ; \ C = \{Y \le c\} \ ; \ D = \{Y \le d\} \ ;$$

for $a < b$ ; $c < d$.




$P(a < X \le b \cap c < Y \le d)$
$= P[(B - A) \cap (D - C)]$
$= P[B \cap (D - C) - A \cap (D - C)]$ ...(*)

(By distributive property of sets)

We know that if $E \subset F \Rightarrow E \cap F = E$, then
$$P(F - E) = P(\bar{E} \cap F) = P(F) - P(E \cap F) = P(F) - P(E) \quad ...(**)$$
Obviously $A \subset B \Rightarrow [A \cap (D - C)] \subset [B \cap (D - C)]$

Hence using (**), we get from (*)
$$P(a < X \le b \cap c < Y \le d) = P[B \cap (D - C)] - P[A \cap (D - C)]$$
$$= P[(B \cap D) - (B \cap C)] - P[(A \cap D) - (A \cap C)]$$

$$= P(B \cap D) - P(B \cap C) - P(A \cap D) + P(A \cap C) \quad \ldots (***)$$

[On using (**), since $C \subset D \Rightarrow (B \cap C) \subset (B \cap D)$ and $(A \cap C) \subset (A \cap D)$]

We have:

$P(B \cap D) = P[X \le b \cap Y \le d] = F(b, d)$.

Similarly

$P(B \cap C) = F(b, c)$ ; $P(A \cap D) = F(a, d)$ and $P(A \cap C) = F(a, c)$

Substituting in (***), we get:

$$P(a < X \le b \cap c < Y \le d) = F(b, d) - F(b, c) - F(a, d) + F(a, c) \quad \ldots (1)$$

*We are given* $\left.\begin{aligned} F(x, y) &= 1, \quad \text{for } x + 2y \ge 1 \\ &= 0, \quad \text{for } x + 2y < 1 \end{aligned}\right\} \quad \ldots (2)$

In (1) let us take : $a = 0$, $b = 1/2$, ; $c = 1/4$, $d = 3/4$ s.t. $a < b$ and $c < d$. Then using (2) we get:

$F(b, d) = 1$ ; $F(b, c) = 1$ ; $F(a, d) = 1$ ; $F(a, c) = 0$.

Substituting in (1) we get:

$P(a < X \le b \cap c < Y \le d) = 1 - 1 - 1 + 0 = -1$ ;

which is not possible since $P(\,.\,) \ge 0$.

Hence $F(x, y)$ defined in (2) cannot be the distribution function of variates $X$ and $Y$.

*(ii)* Let us define the events : $A = \{X \le x\}$ ; $B = \{Y \le y\}$

$\left.\begin{aligned} &\text{Then } P(A) = P(X \le x) = F_X(x) ; \; P(B) = P(Y \le y) = F_Y(y) \\ &\text{and } P(A \cap B) = P(X \le x \cap Y \le y) = F_{XY}(x, y) \end{aligned}\right\} \ldots (3)$

$(A \cap B) \subset A \Rightarrow P(A \cap B) \le P(A) \Rightarrow F_{XY}(x, y) \le F_X(x)$

$(A \cap B) \subset B \Rightarrow P(A \cap B) \le P(B) \Rightarrow F_{XY}(x, y) \le F_Y(y)$

Multiplying these inequalities we get:

$$F^2_{X,Y}(x, y) \le F_X(x) F_Y(y) \Rightarrow F_{XY}(x, y) \le \sqrt{F_X(x) F_Y(y)} \quad \ldots (4)$$

Also $P(A \cup B) \le 1 \Rightarrow P(A) + P(B) - P(A \cap B) \le 1$

$\Rightarrow P(A) + P(B) - 1 \le P(A \cap B)$

$$\Rightarrow F_X(x) + F_Y(y) - 1 \le F_{XY}(x, y) \quad \ldots (5)$$

From (4) and (5) we get:

$F_X(x) + F_Y(y) - 1 \le F_{XY}(x, y) \le \sqrt{F_X(x) F_Y(y)}$, as required.

**Example 5·30.** *If $X$ and $Y$ are two random variables having joint density function*

$$f(x, y) = \frac{1}{8}(6 - x - y) \; ; \; 0 < x < 2, \; 2 < y < 4$$
$$= 0, \quad \textit{otherwise}$$

*Find (i) $P(X < 1 \cap Y < 3)$, (ii) $P(X + Y < 3)$ and (iii) $P(X < 1 \mid Y < 3)$*

**(Madras Univ. B.Sc., Nov. 1986)**

**Solution.** We have

*(i)* $$P(X<1 \cap Y<3)=\int_{-\infty}^{1}\int_{-\infty}^{3} f(x,y)\,dx\,dy$$

$$=\int_0^1\int_2^3 \frac{1}{8}(6-x-y)\,dx\,dy=\frac{3}{8}$$

*(ii)* The probability that $X+Y$ will be less than 3 is

$$P(X+Y<3)=\int_0^1\int_2^{3-x}\frac{1}{8}(6-x-y)\,dx\,dy=\frac{5}{24}$$

*(iii)* The probability that $X<1$ when it is known that $Y<3$ is

$$P(X<1\mid Y<3)=\frac{P(X<1\cap Y<3)}{P(Y<3)}=\frac{3/8}{5/8}=\frac{3}{5}$$

$$\left[P(Y<3)=\int_0^2\int_2^3\frac{1}{8}(6-x-y)\,dx\,dy=\frac{5}{8}\right]$$

**Example 5·31.** *If the joint distribution function of X and Y is given by :*

$$F(x,y)=1-e^{-x}-e^{-y}+e^{-(x+y)};\quad x>0,\ y>0$$
$$=0;\qquad \textit{elsewhere}$$

*(a) Find the marginal densities of X and Y.*

*(b) Are X and Y independent ?*

*(c) Find* $P(X\le 1\cap Y\le 1)$ *and* $P(X+Y\le 1)$. **(I.C.S., 1989)**

**Solution.** *(a)* & *(b)* The joint p.d.f. of the r.v.'s $(X, Y)$ is given by :

$$f_{XY}(x,y)=\frac{\partial^2 F(x,y)}{\partial x\,\partial y}=\frac{\partial}{\partial x}\left[e^{-y}-e^{-(x+y)}\right]$$
$$=e^{-(x+y)};\quad x\ge 0,\ y\ge 0$$
$$=0;\quad \text{otherwise} \qquad \text{...}(i)$$

We have

$$f_{XY}(x,y)=e^{-x}.e^{-y}=f_X(x)f_Y(y) \qquad \text{...}(ii)$$

where $f_X(x)=e^{-x};\ x\ge 0;\qquad f_Y(y)=e^{-y};\ y\ge 0$ ...*(iii)*

*(ii)* $\Rightarrow$ $X$ and $Y$ are independent,

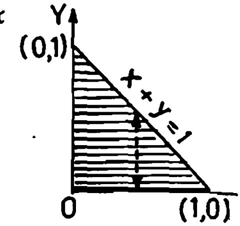
and *(iii)* gives the marginal p.d.f.'s of $X$ and $Y$.

*(c)* $$P(X\le 1\cap Y\le 1)=\int_0^1\int_0^1 f(x,y)\,dx\,dy$$

$$=\left(\int_0^1 e^{-x}\,dx\right)\left(\int_0^1 e^{-y}\,dy\right)$$

$$=\left(1-e^{-1}\right)^2$$

$$P(X+Y \le 1) = \iint\limits_{x+y \le 1} f(x,y) = \int_0^1 \left( \int_0^{1-x} f(x,y)\, dy \right) dx$$




$$= \int_0^1 \left[ e^{-x} \int_0^{1-x} e^{-y}\, dy \right] dx$$

$$= \int_0^1 e^{-x} \left( 1 - e^{-(1-x)} \right) dx = 1 - 2e^{-1}$$

**Example 5·32.** *Joint distribution of X and Y is given by*

$$f(x, y) = 4xy\, e^{-(x^2+y^2)}; \quad x \ge 0,\ y \ge 0.$$

*Test whether X and Y are independent.*

*For the above joint distribution, find the conditional density of X given $Y = y$.* **(Calicut Univ. B.Sc., 1986)**

**Solution.** Joint p.d.f. of $X$ and $Y$ is

$$f(x,y) = 4xy\, e^{-(x^2+y^2)}; \quad x \ge 0,\ y \ge 0.$$

Marginal density of $X$ is given by

$$f_1(x) = \int_0^\infty f(x,y)\, dy = \int_0^\infty 4xy\, e^{-(x^2+y^2)}\, dy$$

$$= 4x\, e^{-x^2} \int_0^\infty y\, e^{-y^2}\, dy$$

$$= 4x\, e^{-x^2} \cdot \int_0^\infty e^{-t} \cdot \frac{dt}{2} \qquad \left( Put\ y^2 = t \right)$$

$$= 2x \cdot e^{-x^2} \left| -e^{-t} \right|_0^\infty$$

$$\Rightarrow \quad f_1(x) = 2x\, e^{-x^2}; \quad x \ge 0$$

Similarly, the marginal p.d.f. of $Y$ is given by

$$f_2(y) = \int_0^\infty f(x,y)\, dx = 2y\, e^{-y^2}; \quad y \ge 0$$

Since $f(x,y) = f_1(x) \cdot f_2(y)$, $X$ and $Y$ are independently distributed. The conditional distribution of $X$ for given $Y$ is given by :

$$f(X=x \mid Y=y)=\frac{f(x,y)}{f_2(y)}$$
$$= 2x\, e^{-x^2}\,;\; x \geq 0\,.$$

## EXERCISE 5(e)

1. (a) Two fair dice are tossed simultaneously. Let $X$ denote the number on the first die and $Y$ denote the number on the second die.

*(i)* Write down the sample space of this experiment.

*(ii)* Find the following probabilities :

*(1)* $P(X+Y=8)$, *(2)* $P(X+Y \geq 8)$, *(3)* $P(X=Y)$,
*(4)* $P(X+Y=6 \mid Y=4)$, *(5)* $P(X-Y=2)$.

**(Sardar Patel Univ. B.Sc., 1991)**

2. *(a)* Explain the concepts *(i)* conditional probability, *(ii)* random variable, *(iii)* independence of random variables, and *(iv)* marginal and conditional probability distributions.

*(b)* Explain the notion of the joint distribution of two random variables. If $F(x, y)$ be the joint distribution function of $X$ and $Y$, what will be the distribution functions for *the marginal distribution of X and Y ?*

What is meant by the *conditional distribution* of $Y$ under the condition that $X = x$ ? Consider separately the cases where *(i)* $X$ and $Y$ are both discrete and *(ii)* $X$ and $Y$ are both continuous.

3. The joint probability distribution of a pair of random variables is given by the following table :-

| Y \ X | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 0.1 | 0.1 | 0.2 |
| 2 | 0.2 | 0.3 | 0.1 |

Find :

*(i)* The marginal distributions.

*(ii)* The conditional distribution of $X$ given $Y=1$.

*(iii)* $P\{(X+Y)<4\}$.

4. *(a)* What do you mean by marginal and conditional distributions ? The following table represents the joint probability distribution of the discrete random variable $(X, Y)$

| Y \ X | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 1/12 | 1/6 | 0 |
| 2 | 0 | 1/9 | 1/5 |
| 3 | 1/18 | 1/4 | 2/15 |

*(i)* Evaluate marginal distribution of $X$.

*(ii)* Evaluate the conditional distribution of Y given $X = 2$.

**(Aligarh Univ. B.Sc., 1992)**

*(b)* Two discrete random variables $X$ and $Y$ have

$$P(X=0,\ Y=0)=\frac{2}{9}\ ;\ P(X=0,\ Y=1)=\frac{1}{9}$$

$$P(X=1,\ Y=0)=\frac{1}{9}\ ;\ P(X=1,\ Y=1)=\frac{5}{9}$$

Examine whether $X$ and $Y$ are independent.

**(Kerala Univ. B.Sc., Oct. 1987)**

5. *(a)* Let the joint p.m.f. of $X_1$ and $X_2$ be

$$p(x_1, x_2)=\frac{x_1+x_2}{21}\ ;\ x_1=1,2,3\ ;\ x_2=1,2$$
$$=0\text{, otherwise}$$

Show that marginal p.m.f.'s of $X_1$ and $X_2$ are

$$p_1(x_1)=\frac{2x_1+3}{21}\ ;\ x_1=1,2,3\ ;\qquad p_2(x_2)=\frac{6+3x_2}{21}\ ;\ x_2=1,2$$

*(b)* Let

$$f(x_1, x_2)=C(x_1x_2+e^{x_1})\ ;\ 0<(x_1,\ x_2)<1$$
$$=0\text{, elsewhere}$$

*(i)* Determine $C$.

*(ii)* Examine whether $X_1$ and $X_2$ are stochastically independent.

**Ans.** *(i)* $C=\dfrac{4}{4e-3}$, *(ii)* $g(x_1)=C\left(\frac{1}{2}x_1+e^{x_1}\right)$, $g(x_2)=C\left(\frac{1}{2}x_2+e-1\right)$

Since $g(x_1)\cdot g(x_2)\neq f(x_1, x_2)$, $X_1$ and $X_2$ are not stochastically independent.

6. Find $k$ so that $f(x, y)=kxy$, $1\le x\le y\le 2$ will be a probability density function. **(Mysore Univ. B.Sc., 1986)**

**Hint.** $\displaystyle\iint f(x,y)\,dx\,dy=1\ \Rightarrow\ k\int_1^2 x\left(\int_x^2 y\,dy\right)dx=1\ \Rightarrow\ k=8/9$

7. *(a)* If

$$f(x, y)=e^{-(x+y)}\ ;\ x\ge 0,\ y\ge 0$$
$$=0\text{, elsewhere}$$

is the joint probability density function of random variables $X$ and $Y$, find *(i)* $P(X<1)$, *(ii)* $P(X>Y)$, and *(iii)* $P(X+Y<1)$.

**Ans.** *(i)* $1-\dfrac{1}{e}$, *(ii)* $\dfrac{1}{2}$ and *(iii)* $1-\dfrac{2}{e}$

*(b)* The joint frequency function of $(X, Y)$ is given to be

$$f(x, y) = A e^{-x-y} \; ; \quad 0 \le x \le y, \; 0 \le y < +\infty$$
$$= 0 \quad ; \quad \text{otherwise}$$

*(i)* Determine $A$.
*(ii)* Find the marginal density function of $X$.
*(iii)* Find the marginal density function of $Y$.
*(iv)* Examine if $X$ and $Y$ are independent.
*(v)* Find the conditional density function of $Y$ given $X = 2$.

**[Madras Univ. B.Sc. (Main Stat.), 1992]**

*(c)* Suppose that the random variables $X$ and $Y$ have the joint p.d.f.

$$f(x, y) = \begin{cases} kx(x-y), & 0 < x < 2, \; -x < y < x \\ 0, & \textit{elsewhere}. \end{cases}$$

*(i)* Evaluate the constant k.
*(ii)* Find the marginal probability density functions of the random variables.

**(South Gujarat Univ. B.Sc., 1988)**

8. *(a)* Two-dimensional random variable $(X, Y)$ have the joint density

$$f(x, y) = 8xy, \; 0 < x < y < 1$$
$$= 0, \; \text{otherwise}$$

*(i)* Find. $P(X < 1/2 \cap Y < 1/4)$.
*(ii)* Find the marginal and conditional distributions.
*(iii)* Are $X$ and $Y$ independent? Give reasons for your answer.

**(South Gujarat Univ. B.Sc., 1992)**

Ans. $f_1(x) = 4x(1-x^2), 0 < x < 1$; $= 0$, *otherwise*; $f_2(y) = 4y^3, 0 < y < 1$ | $f_1(x \mid y) = 2x/y^2 \;:\; 0 < x < y, 0 < y < 1$; $f_2(y \mid x) = 2y/(1-x^2)\,;\; x < y < 1, 0 < x < 1$

9. *(a)* The random variables $X$ and $Y$ have the joint density function :

$$f(x, y) = 2, \text{ if } x + y \le 1, \; x \ge 0 \text{ and } y \ge 0$$
$$= 0, \; \textit{otherwise}$$

Find the conditional distribution of $Y$, given $X = x$.

**(Calcutta Univ. B.Sc. (Hons.), 1984)**

*(b)* The random variables $X$ and $Y$ have the joint distribution given by the probability density function :

$$f(x, y) = \begin{cases} 6(1-x-y), & \text{for } x > 0, \; y > 0, \; x + y < 1 \\ 0, & \textit{elsewhere} \end{cases}$$

Find the marginal distributions of $X$ and $Y$. Hence examine if $X$ and $Y$ are independent. **[Calcutta Univ. B.Sc. (Hons.), 1986)**

10. If the joint distribution function of $X$ and $Y$ is given by

$$F(x, y) = (1 - e^{-x})(1 - e^{-y}) \quad \text{for } x > 0, \; y > 0$$
$$= 0, \; \text{elsewhere}$$

Find $P(1 < X < 3,\ 1 < Y < 2)$. [Delhi Univ. M.A.(Econ.), 1988]

**Hint.** Reqd. Prob. $= \left(\int_1^3 e^{-x}\, dx\right)\left(\int_1^2 e^{-y}\, dy\right) = \left(1 - e^{-3}\right)\left(1 - e^{-2}\right)$

**11.** Let $X$ and $Y$ be two random variables with the joint probability density function

$$f(x,y) = \begin{cases} 8xy, & 0 < x \le y < 1 \\ 0 & , \textit{otherwise} \end{cases}$$

Obtain :

*(i)* the joint distribution function of $X$ and $Y$.

*(ii)* the marginal probability density function of $Y$; and

*(iii)* $P\left(X \le \frac{1}{4} \mid \frac{1}{2} < Y \le 1\right)$.

**12.** Let $X$ and $Y$ be jointly distributed with p.d.f.

$$f(x,y) = \begin{cases} \frac{1}{4}(1 + xy), & |x| < 1,\ |y| < 1 \\ 0 & , \textit{otherwise} \end{cases}$$

Show that $X$ and $Y$ are not independent but $X^2$ and $Y^2$ are independent.

**Hint.** $f_1(x) = \int_{-1}^{1} f(x,y)\, dy = \frac{1}{2},\ -1 < x < 1;$

$$f_2(y) = \int_{-1}^{1} f(x,y)\, dx = \frac{1}{2},\ -1 < y < 1$$

Since $f(x,y) \ne f_1(x) f_2(y)$, $X$ and $Y$ are not independent. However,

$$P(X^2 \le x) = P(|X| \le \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} f_1(x)\, dx = \sqrt{x}$$

$$\begin{aligned} P(X^2 \le x \cap Y^2 \le y) &= P(|X| \le \sqrt{x} \cap |Y| \le \sqrt{y}) \\ &= \int_{-\sqrt{x}}^{\sqrt{x}} \left[\int_{-\sqrt{y}}^{\sqrt{y}} f(u,v)\, dv\right] du \\ &= \sqrt{x} \cdot \sqrt{y} \\ &= P(X^2 \le x) \cdot P(Y^2 \le y) \end{aligned}$$

$\Rightarrow$ $X^2$ and $Y^2$ are independent.

**13.** *(a)* The joint probability density function of the two dimensional random variable $(X, Y)$ is given by :

$$f(x,y)=\begin{cases} x^3y^3/16\ , & 0\le x\le 2\ ,\ 0\le y\le 2 \\ 0 & ,\ elsewhere \end{cases}$$

Find the marginal densities of $X$ and $Y$. Also find the cumulative distribution functions for $X$ and $Y$. **(Annamalai Univ. B.E., 1986)**

**Ans.** $f_X(x)=\frac{x^3}{4}\ ;\ 0\le x\le 2\ ;\quad f_Y(y)=\frac{y^3}{4}\ ;\ 0\le y\le 2$

$$F_X(x)=\begin{cases} 0 & ;\ x<o \\ x^4/16 & ;\ 0\le x\le 2 \\ 1 & ;\ x>2 \end{cases}\qquad F_Y(y)=\begin{cases} 0 & ;\ y<0 \\ y^4/16 & ;\ 0\le y\le 2 \\ 1 & ;\ y>2 \end{cases}$$

*(b)* The joint probability density function of the two dimensional random variable $(X, Y)$ is given by :

$$f(x,y)=\begin{cases} \frac{8}{9}xy & ,\ 1\le x\le y\le 2 \\ 0 & ,\ elsewhere \end{cases}$$

*(i)* Find the marginal density functions of $X$ and $Y$,

*(ii)* Find the conditional density function of $Y$ given $X = x$, and conditional density funciton of $X$ given $Y = y$.

**[Madras Univ. B.Sc. (Stat. Main), 1987]**

**Ans.** *(i)* $f_X(x)=\int_x^2 f(x,y)\,dy=\frac{4}{9}x\left(4-x^2\right)\ ;\ 1\le x\le 2$

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

$$f_Y(y)=\int_1^y f(x,y)\,dx=\frac{4}{9}y\left(y^2-1\right)\ ;\ 1\le y\le 2$$

$$f_{X|Y}(x\mid y)=\frac{2x}{y^2-1}\qquad ;\ 1\le x\le y$$

$$f_{Y|X}(y\mid x)=\frac{f(x,y)}{f_X(y)}=\frac{2y}{4-x^2}\ ;\ x\le y\le 2$$

14. The two random variables $X$ and $Y$ have, for $X = x$ and $Y = y$, the joint probability density function :

$$f(x,y)=\frac{1}{2x^2y}\ ,\quad \text{for } 1\le x<\infty \ \text{ and }\ \frac{1}{x}<y<x$$

Derive the marginal distributions of $X$ and $Y$. Further obtain the conditional distribution of $Y$ for $X = x$ and also that of $X$ given $Y = y$.

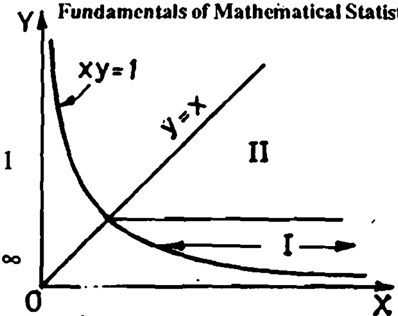
**(Civil Services Main, 1986)**

**Hint.** $f_X(x)=\int_y f(x,y)\,dy=\int_{1/x}^{x} f(x,y)\,dy$

$$f_Y(y) = \int_x f(x,y)\,dx$$

$$= \int_{1/y}^{\infty} f(x,y)\,dx;\quad 0 \le y \le 1$$

$$= \int_{y}^{\infty} f(x,y)\,dx;\quad 1 \le y < \infty$$




15. Show that the conditions for the function

$$f(x,y) = k\,\exp\left[A x^2 + 2\,H x y + B y^2\right],\ -\infty < (x, y) < \infty$$

to be a bivariate p.d.f. are

*(i)* $A \le 0$, *(ii)* $B \le 0$ *(iii)* $A B - H^2 \ge 0$.

Further show that under these conditions

$$k = \frac{1}{\pi}\left(A B - H^2\right)^{1/2}$$

**Hint.** $f(x,y)$ will represent the p.d.f. of a bivariate distriibution if and only if

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\ dx\ dy = 1$$

$$\Rightarrow\quad k\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left[A x^2 + 2 H x y + B y^2\right] dx\ dy = 1 \qquad \ldots(*)$$

We have

$$A x^2 + 2 H x y + B y^2 = A\left[x^2 + \frac{2H}{A} x y + \frac{B}{A} y^2\right]$$

$$= A\left[\left(x + \frac{H}{A} y\right)^2 + \frac{AB - H^2}{A^2}\,.\,y^2\right] \qquad \ldots(**)$$

Similarly, we can write

$$A x^2 + 2 H x y + B y^2 = B\left[\left(y + \frac{H}{B} x\right)^2 + \frac{AB - H^2}{B^2} x^2\right] \qquad \ldots(***)$$

Substituting from (**) and (***) in (*) we observe that the double integral on the left hand side will converge if and only if

$$A \le 0,\ B \le 0 \text{ and } AB - H^2 \ge 0,$$

as desired.

Let us take $A = -a$ ; $B = -b$ ; $H = h$ so that $AB - H^2 = ab - h^2$, where $a > 0, b > 0$.

Substituting in (*), we get

$$k\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left[-\frac{ab-h^2}{a}y^2-\frac{1}{a}(-ax+hy)^2\right]dx\,dy=1$$

$$\Rightarrow \quad k\int_{-\infty}^{\infty}\left[\exp\left(-\frac{ab-h^2}{a}y^2\right)\cdot\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{a}(ax-hy)^2\right\}dx\right]dy$$

$$=1 \qquad \ldots (****)$$

(By Fubini's theorem)

Now $\displaystyle\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{a}(ax-hy)^2\right\}dx=\int_{-\infty}^{\infty}\exp\left(-\frac{u^2}{a}\right)\frac{du}{a}$

$(ax-hy=u)$

$$=\frac{1}{a}\sqrt{\pi}\sqrt{a}=\sqrt{\frac{\pi}{a}}$$

$$\left(\because \int_{-\infty}^{\infty}e^{-c^2u^2}\,du=\frac{\sqrt{\pi}}{c}\right)$$

Hence from (****), we get

$$k\sqrt{\frac{\pi}{a}}\int_{-\infty}^{\infty}\exp\left\{-\frac{ab-h^2}{a}y^2\right\}dy=1$$

$$\Rightarrow \quad k\sqrt{\frac{\pi}{a}}\cdot\sqrt{\frac{\pi a}{ab-h^2}}=1$$

$$\Rightarrow \quad k=\frac{1}{\pi}\sqrt{ab-h^2}=\frac{1}{\pi}\sqrt{AB-H^2}.$$

## OBJECTIVE TYPE QUESTIONS

I. Which of the following statements are TRUE or FALSE.

*(i)* Given a continuous random variable $X$ with probability density function $f(x)$, then $f(x)$ cannot exceed unity.

*(ii)* A random variable $X$ has the following probability density function :

$$f(x)=x,\ 0<x<1$$
$$=0, \text{ elsewhere}$$

*(iii)* The function defined as

$$f(x)=|x|,\ -1<x<1$$
$$=0, \text{ elsewhere}$$

is a possible probability density function.

*(iv)* The following represents joint probability distribution.

| Y \ X | 1 | 2 | 3 |
|---|---|---|---|
| –1 | 1/9 | 1/18 | 1/18 |
| 0 | 1/18 | 2/9 | 3/9 |
| 1 | 1/8 | 1/18 | 1/18 |

**II.** Fill in the blanks :

*(i)* If $p_1(x)$ and $p_2(y)$ be the marginal probability functions of two independent discrete random variablies $X$ and $Y$, then their joint probability function

$$p(x, y) = \ldots$$

*(ii)* The function $f(x)$ defined as

$$f(x) = |x|, \; -1 < x < 1$$
$$= 0, \text{ elsewhere}$$

is a possible ......

**5·6. Transformation of One-dimensional Random Variable.** Let $X$ be a random variable defined on the event space $S$ and let $g(\cdot)$ be a function such that $Y = g(X)$ is also a r.v. defined on $S$. In this section we shall deal with the following problem :

"Given the probability density of a r.v. $X$, to determine the density of a new r.v. $Y = g(X)$."

It can be proved in general that, if $g(\cdot)$ is any continuous function, then the distribution of $Y = g(X)$ is uniquely determined by that of $X$. The proof of this result is rather difficult and beyond the scope of this book. Here we shall consider the following, relatively simple theorem.

**Theorem 5·9.** *Let $X$ be a continuous r.v. with p.d.f. $f_X(x)$. Let $y = g(x)$ be strictly monotonic (increasing or decreasing) function of $x$. Assume that $g(x)$ is differentiable (and hence continuous) for all $x$. Then the p.d.f. of the r.v. $Y$ is given by*

$$h_Y(y) = f_X(x) \left| \frac{dx}{dy} \right|,$$

*where $x$ is expressed in terms of $y$.*

**Proof.** Case (i). $y = g(x)$ *is strictly increasing function of $x$ (i.e., $dy/dx > 0$*. The *d.f.* of $Y$ is given by

$$H_Y(y) = P(Y \leq y) = P[g(X) \leq y] = P(X \leq g^{-1}(y))],$$

the inverse exists and is unique, since $g(\cdot)$ is strictly increasing.

$$\therefore \qquad H_Y(y) = F_X\left[g^{-1}(y)\right], \text{ where } F \text{ is the d.f. of } X$$

$$= F_X(x) \qquad \left[\because y = g(x) \Rightarrow g^{-1}(y) = x\right]$$

Differentiating w.r.t. $y$, we get

$$h_Y(y) = \frac{d}{dy}[F_X(x)] = \frac{d}{dx}\left(F_X(x)\right)\frac{dx}{dy}$$

$$= f_X(x)\frac{dx}{dy} \qquad \ldots (*)$$

Case (ii). $y = g(x)$ *is strictly monotonic decreasing.*

$$H_Y(y) = P(Y \leq y) = P[g(X) \leq y] = P[X \geq g^{-1}(y)]$$
$$= 1 - P\left[X \leq g^{-1}(y)\right] = 1 - F_X\left[g^{-1}(y)\right] = 1 - F_X(x),$$

where $x = g^{-1}(y)$, the inverse exists and is unique. Differentiating w.r.t. $y$, we get

$$h_Y(y) = \frac{d}{dx}\left[1 - F_X(x)\right]\frac{dx}{dy} = -f_X(x) \cdot \frac{dx}{dy}$$

$$= f_X(x) \cdot \frac{-dx}{dy} \qquad \text{...}(**)$$

Note that the algebraic sign (–ive) obtained in (**) is correct, since $y$ is a decreasing function of $x \Rightarrow x$ is a decreasing function of $y \Rightarrow dx/dy < 0$.

The results in (*) and (**) can be combined to give

$$h_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

**Example 5·33.** *If the cumulative distribution function of X is $F(x)$, find the cumulative distribution function of*

*(i)* $Y = X + a$, *(ii)* $Y = X - b$, *(iii)* $Y = aX$,

*(iv)* $Y = X^3$, and *(v)* $Y = X^2$

*What are the corresponding probability density functions ?*

**Solution.** Let $G(\cdot)$ be the c.d.f. of $Y$. Then

*(i)* $G(x) = P(Y \leq x) = P[X + a \leq x] = P[X \leq x - a] = F(x - a)$

*(ii)* $G(x) = P(Y \leq x) = P[X - b \leq x] = P[X \leq x + b] = F(x + b)$

*(iii)* $G(x) = P[aX \leq x] = P\left[X \leq \frac{x}{a}\right], \; a > 0$

$$= F\left(\frac{x}{a}\right), \text{ if } a > 0$$

and $G(x) = P\left[X \geq \frac{x}{a}\right] = 1 - P\left[X < \frac{x}{a}\right]$

$$= 1 - F\left(\frac{x}{a}\right), \text{ if } a < 0$$

*(iv)* $G(x) = P[Y \leq x] = P[X^3 \leq x] = P\left[X \leq x^{1/3}\right] = F\left(x^{1/3}\right)$

*(v)* $G(x) = P\left[X^2 \leq x\right] = \left[-x^{1/2} \leq X \leq x^{1/2}\right]$

$$= P\left[X \leq x^{1/2}\right] - P\left[X \leq -x^{1/2}\right]$$

$= 0, \qquad \text{if } x < 0$

$= F(\sqrt{x}) - F(-\sqrt{x} - 0), \qquad \text{if } x > 0$

| Variable | d.f. | p.d.f. |
|---|---|---|
| $X$ | $F(x)$ | $f(x)$ |
| $X - a$ | $F(x+a)$ | $f(x+a)$ |
| $aX$ | $\left.\begin{array}{l} F(x/a) \quad a > 0 \\ 1 - F(x/a), a < 0 \end{array}\right\}$ | $(1/a)\, f(x/a), \ a > 0$ <br> $(-1/a)\, f(x/a), \ a < 0$ |
| $X^2$ | $\left.\begin{array}{l} F(\sqrt{x}) - F(-\sqrt{x} - 0) \\ \quad \text{for } x > 0 \\ 0, \textit{ otherwise} \end{array}\right\}$ | $\left.\begin{array}{l} \dfrac{1}{2(\sqrt{x})} [f\sqrt{x} + f(-\sqrt{x})] \\ \quad \text{for } x > 0 \\ = 0 \text{ for } x \le 0 \end{array}\right\}$ |
| $X^3$ | $F(x^{1/3})$ | $\frac{1}{3} f(x^{1/3}) \cdot \frac{1}{x^{2/3}}$ |

## EXERCISE 5(f)

1. *(a)* A random variable $X$ has $F(x)$ as its distribution function [$f(x)$ is the density function ]. Find the distribution and the density functions of the random variable :

*(i)* $Y = a + bX$, $a$ and $b$ are real numbers, *(ii)* $Y = X^{-1}$, $[P(X = 0) = 0]$, *(iii)* $Y = \tan X$, and *(iv)* $Y = \cos X$.

*(b)* Let $f(x) = \begin{cases} ½, & -1 < x < 1 \\ 0, & \textit{elsewhere} \end{cases}$

be the p.d.f. of the r.v. $X$. Find the distribution function and the p.d.f. of $Y = X^2$.

**[ Delhi Univ. B.Sc. (Maths Hons.), 1988 ]**

**Hint.** $F(x) = P(X \le x) = \int_{-1}^{x} f(x)\, dx = \frac{1}{2}(x + 1)$ ...(*)

Distribution function $G(\cdot)$ of $Y = X^2$ is given by :

$G_Y(x) = F(\sqrt{x}) - F(-\sqrt{x}) \quad ; \quad x > 0$ [ c.f. Example 5·33 *(v)* ]

$= \frac{1}{2}(\sqrt{x} + 1) - \frac{1}{2}(-\sqrt{x} + 1)$ [ From (*) ]

$= \sqrt{x} \quad ; \quad 0 < x < 1$

( As $-1 < x < 1$, $Y = X^2$ lies between 0 and 1 )

p.d.f. of $Y = X^2$ is $g(x) = G'(x) = \dfrac{1}{2\sqrt{x}} \quad ; \quad 0 < x < 1$

2. Let $X$ be a continuous random variable with p.d.f. $f(x)$. Let $Y = X^2$. Show that the random variable $Y$ has p.d.f. given by

$$g(y) = \begin{cases} \frac{1}{2\sqrt{y}} [f(\sqrt{y}) + f(-\sqrt{y})], & y > 0 \\ 0, & y \le 0 \end{cases}$$

3. Find the distribution and densitiy functions for *(i)* $Y = aX + b$, $a \neq 0$, $b$ real, *(ii)* $Y = e^X$, assuming that $F(x)$ and $f(x)$, the distribution and the density of $X$ are known.

**Ans.** *(i)* $\left.\begin{array}{ll} G(y) = F[(y-b)/a], & \text{if } a > 0 \\ G(y) = 1 - F[(y-b)/a], & \text{if } a < 0 \end{array}\right\}$ $g_1(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right)$

*(ii)* $\left.\begin{array}{ll} G(y) = F(\log y), & y > 0 \\ \quad = 0, & y \le 0 \end{array}\right\}$ $\begin{array}{ll} g(y) = \frac{1}{y} f(\log y), & y > 0 \\ \quad = 0, & y \le 0 \end{array}$

4. *(a)* The random variable $X$ has an exponential distribution

$$f(x) = e^{-x}, \quad 0 < x \le \infty$$

Find the density function of the variable *(i)* $Y = 3X + 5$, *(ii)* $Y = X^3$.

*(b)* Suppose that $X$ has p.d.f.,

$$\begin{aligned} f(x) &= 2x, \quad 0 < x < 1 \\ &= 0, \quad \textit{elsewhere} \end{aligned}$$

Find the p.d.f. of $Y = 3X + 1$.

**Ans.** $g(y) = \frac{2}{9}(y-1)$, $1 < y < 4$

5. Let $X$ be a random variable with p.d.f.

$$\begin{aligned} f(x) &= \tfrac{2}{9}(x+1) \qquad -1 < x < 2 \\ &= 0, \quad \textit{elsewhere} \end{aligned}$$

Find the p.d.f. of $U = X^2$. **[ Poona Univ. B.E., 1992 ]**

6. Let the p.d.f. of $X$ be

$$\begin{aligned} f(x) &= \tfrac{1}{6}, \quad -3 \le x \le 3 \\ &= 0, \quad \textit{elsewhere} \end{aligned}$$

Find the p.d.f. of $Y = 2X^2 - 3$.

7. Let $X$ be a random variable with the distribution function :

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$$

Determine the distribution function $F_Y(y)$ of the random variable $Y = \sqrt{X}$ and hence compute mean of $Y$. **[ Calcutta Univ. B.A.(Hons.), 1986 ]**

**5·7. Transformation of Two-dimensional Random Variable.** In this section we shall consider the problem of change of variables in the two-dimensional case. Let the *r.v.'s* $U$ and $V$ by the transformation $u = u(x, y)$, $v = v(x, y)$, where $u$ and $v$ are continuously differentiable functions for which Jacobian of transformation

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix}$$

is either $> 0$ or $< 0$ throughout the $(x, y)$ plane so that the inverse transformation is uniquiely given by $x = x(u, v)$, $y = y(u, v)$.

**Theorem 5·10.** *The joint p.d.f.* $g_{UV}(u, v)$ *of the transformed variables* $U$ *and* $V$ *is given by*

$$g_{UV}(u, v) = f_{XY}(x, y) \,.\, |J|$$

*where* $|J|$ *is the modulus value of the Jacobian of transformation and* $f(x, y)$ *is expressed in terms of* $u$ *and* $v$.

**Proof.** $P(x < X \leq x + dx,\ y < Y \leq y + dy)$

$$= P(u < U \leq u + du,\ v < V \leq v + dv)$$

$$\Rightarrow \quad f_{XY}(x, y)\ dx\ dy = g_{UV}(u, v)\ du\ dv$$

$$\Rightarrow \quad g_{UV}(u, v)\ du\ dv = f_{XY}(x, y) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du\ dv$$

$$\Rightarrow \quad g_{UV}(u, v) = f_{XY}(x, y) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| = f_{XY}(x, y)\ |J|$$

**Theorem 5·11.** *If* $X$ *and* $Y$ *are independent continuous r.v.'s, then the p.d.f. of* $U = X + Y$ *is given by*

$$h(u) = \int_{-\infty}^{\infty} f_X(v) f_Y(u - v)\ dv$$

**Proof.** Let $f_{XY}(x, y)$ be the joint *p.d.f.* of independent continuous *r.v.'s* $X$ and $Y$ and let us make the transformation :

$$u = x + y,\ v = x \quad \Rightarrow \quad x = v,\ y = u - v$$

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

Thus the joint p.d.f. of *r.v.'s* $U$ and $V$ is given by

$$g_{UV}(u, v) = f_{XY}(x, y)\ |J|$$

$$= f_X(x) \cdot f_Y(y)\ |J|$$

(Since $X$ and $Y$ are independent)

$$= f_X(v) \cdot f_Y(u - v)$$

The marginal density of $U$ is given by

$$h(u)=\int_{-\infty}^{\infty} g_{UV}(u,v)\ dv$$

$$=\int_{-\infty}^{\infty} f_X(v)\, f_Y(u-v)\ dv$$

**Remark.** The function $h(\cdot)$ is given a special name and is said to be the *convolution* of $f_X(\cdot)$ and $f_Y(\cdot)$ and we write

$$h(\cdot)=f_X(\cdot) * f_Y(\cdot)$$

**Example 5·34.** *Let $(X, Y)$ be a two-dimensional non-negative continuous r.v. having the joint density :*

$$f(x,y)=\begin{cases} 4xy\ e^{-(x^2+y^2)} & ;\ x\geq 0,\ y\geq 0 \\ 0 & ,\ \textit{elsewhere}\end{cases}$$

Prove that the densitiy function of $U=\sqrt{X^2+Y^2}$ is

$$h(u)=\begin{cases} 2u^3\ e^{-u^2} & ,\ 0\leq u<\infty \\ 0 & ,\ \textit{elsewhere}\end{cases}$$

**[ Meerut Univ. M.Sc., 1986 ]**

**Solution.** Let us make the transformation :

$$u=\sqrt{x^2+y^2} \quad \text{and} \quad v=x$$

$\Rightarrow \quad v\geq 0,\ u\geq 0$ and $u\geq v \quad \Rightarrow \quad u\geq 0$ and $0\leq v\leq u$

The Jacobian of transformation $J$ is given by

$$\frac{1}{J}=\frac{\partial(u,v)}{\partial(x,y)}=\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v}\end{vmatrix}=-\cdot\frac{y}{\sqrt{x^2+y^2}}$$

The joint p.d.f. of $U$ and $V$ is given by

$$g(u,v)=f(x,y)\ |J|$$

$$=4xy\ e^{-(x^2+y^2)}\left|-\frac{\sqrt{x^2+y^2}}{y}\right|$$

$$=4x\sqrt{x^2+y^2}\ e^{-(x^2+y^2)}$$

$$=\begin{cases} 4vu\,.\,e^{-u^2} & ;\ u\geq 0,\ 0\leq v\leq u \\ 0\ , & \textit{otherwise}\end{cases}$$

Hence the density function of $U = \sqrt{X^2 + Y^2}$ is

$$h(u) = \int_0^u g(u, v)\, dv = 4u\, e^{-u^2} \int_0^u v\, dv$$

$$= \begin{cases} 2u^3 e^{-u^2}, & u \geq 0 \\ 0, & elsewhere \end{cases}$$

**Example 5·35.** *Let the probability density function of the random variable* $(X, Y)$ *be*

$$f(x, y) = \begin{cases} \alpha^{-2} e^{-(x+y)/\alpha} & ; \quad x, y > 0,\ \alpha > 0 \\ 0 & , \quad elsewhere \end{cases}$$

*Find the distribution of* $\frac{1}{2}(X - Y)$. **[ Nagpur Univ. B.E., 1988 ]**

**Solution.** Let us make the transformation :

$$u = \frac{1}{2}(x - y) \quad \text{and} \quad v = y$$

$$\Rightarrow \qquad x = 2u + v \quad \text{and} \quad y = v$$

The Jacobian of the transformation is :

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} = 2$$

Thus, the joint p.d.f. of the random variables $(U, V)$ is given by :

$$g(u, v) = \begin{cases} \dfrac{2}{\alpha^2} e^{-(2/\alpha)(u+v)}, & -\infty < u < \infty,\ v > -2u, \text{ if } u < 0 \\ & v > 0 \text{ if } u \geq 0 \text{ and } \alpha > 0 \\ 0, & elsewhere \end{cases}$$

The marginal p.d.f. of $U$ is given by

$$g_U(u) = \begin{cases} \displaystyle\int_{-2u}^{\infty} \frac{2}{\alpha^2} \exp\{-(2/\alpha)(u + v)\}\, dv \\ \quad = \dfrac{1}{\alpha} e^{-2u/\alpha}, \quad u < 0 \\ \displaystyle\int_0^{\infty} \frac{2}{\alpha} \exp\{-(2/\alpha)(u + v)\}\, dv \\ \quad = \dfrac{1}{\alpha} e^{-2u/\alpha}, \quad u \geq 0 \end{cases}$$

Hence

$$g_U(u) = \frac{1}{\alpha} e^{-(2/\alpha)|u|} \quad ; \quad -\infty < u < \infty$$

**Example 5·36.** *Given the joint density function of X and Y as*

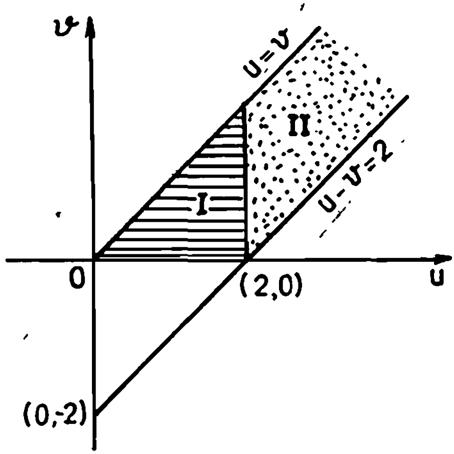
$$f(x,y)=\frac{1}{2}x\,e^{-y}\ ;\ 0<x<2,\ y>0$$
$$=0,\ \textit{elsewhere}$$

*Find the distribution of $X+Y$.*

**Solution.** Let us make the transformation :

$$u=x+y \text{ and } v=y \quad\Rightarrow\quad y=v,\ x=u-v$$

The Jacobian of transformation $J=\dfrac{\partial\,(x,y)}{\partial\,(u,v)}=1$ and the region $0<x<2$ and $y>0$ transforms to $0<u-v<2$ and $v>0$ as shown in the following figure.




The joint density function of $U$ and $V$ is given by

$$g(u,v)=\frac{1}{2}(u-v)\,e^{-v}\ ;\ 0<v<u,\ u>0$$

To find the density of $U=X+Y$, we split the range of U into two parts ($i$) $0<u\le 2$ (region I) ($ii$) $u>2$ (region II) (which is suggested by the diagram).

For $0<u\le 2$, (Region I) :

$$h(u)=\int_0^u g(u,v)\,dv=\frac{1}{2}\int_0^u (u-v)\,e^{-v}\,dv$$

$$=\frac{1}{2}\left|\,-e^{-v}(u-v)+e^{-v}\,\right|_{v=0}^{v=u} \qquad \text{(Integration by parts)}$$

$$=\frac{1}{2}\left(e^{-u}+u-1\right)$$

For $2 < u < \infty$, (Region II):

$$h(u) = \frac{1}{2}\int_{u-2}^{u} (u-v)\, e^{-v}\, dv$$

$$= \frac{1}{2}\left| e^{-v}\,(1+v-u)\right|_{v=u-2}^{v=u}$$

$$= \frac{1}{2}\, e^{-u}\left(1+e^{2}\right) \qquad \text{(on simplification)}$$

Hence

$$g(u) = \begin{cases} \frac{1}{2}\left(e^{-u} + u - 1\right), & 0 < u \leq 2 \\ \frac{1}{2}\, e^{-u}\left(1+e^{2}\right), & 2 < u < \infty \\ 0, & \textit{elsewhere} \end{cases}$$

## MISCELLANEOUS EXERCISE ON CHAPTER FIVE

1. 4 coins are tossed. Let $X$ be the number of heads and $Y$ be the number of heads minus the number of tails. Find the probability function of $X$, the probability function of Y and $P(-2 \leq Y < 4)$.

**Ans.** Probability function of $X$ is

| Values of $X$, $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $p_1(x)$ | $\frac{1}{16}$ | $\frac{4}{16}$ | $\frac{6}{16}$ | $\frac{4}{16}$ | $\frac{1}{16}$ |

Probability function of $Y$ is

| Values of $Y$, $y$ | 4 | 2 | 0 | $-2$ | $-4$ |
|---|---|---|---|---|---|
| $p_2(y)$ | $\frac{1}{16}$ | $\frac{4}{16}$ | $\frac{6}{16}$ | $\frac{4}{16}$ | $\frac{1}{16}$ |

$$P(-2 \leq Y < 4) = \frac{4+6+4}{16} = \frac{7}{8}.$$

2. A random process gives measurements $X$ between 0 and 1 with a probability density function

$$f(x) = 12x^3 - 21x^2 + 10x, \quad 0 \leq x \leq 1$$
$$= 0, \quad \textit{elsewhere}$$

(*i*) Find $P(X \leq \frac{1}{2})$ and $P(X > \frac{1}{2})$

(*ii*) Find a number $k$ such that $P(X \leq k) = \frac{1}{2}$.

**Ans.** (*i*) 9/16, 7/16, (*ii*) $k = 0{\cdot}452$.

3. Show that for the distribution

$$dF = y_0 \left[ 1 - \frac{|x-b|}{a} \right] dx, \quad b-a < x < b+a$$
$$= 0, \ \textit{otherwise},$$
$$y_0 = \frac{1}{a}, \ \textit{mean} = b \text{ and } \textit{variance} = a^2/6$$

4. A ray of light is sent in a random direction towards the $x$-axis from a station $Q$ (0, 1) on the $y$-axis and the ray meets the $x$-axis at a point $P$. Find the probability density function of the abscissa of $P$.

**[Calcutta Univ. B.Sc.(Hons.), 1982]**

5. Let $X$ be a continuous variate with p.d.f.

$$f(x) = k(x - x^2) \ ; \ a < x < b, \ k > 0$$

What are the possible values of $a$ and $b$ and what is $k$ ?

**[Delhi Univ. B.Sc.(Maths Hons.), 1989]**

6. Pareto distribution with parameters $r$ and $A$ is given by the probability density function

$$f(x) = rA^r \frac{1}{x^{r+1}}, \quad \text{for } x \geq A$$
$$= 0, \ x < A, \ r > 0$$

Show that it has a finite $n$th moment if and only if $n < r$. Find the mean and variance of the distribution.

7. For a continuous random variable $X$, defined in the range $(0 \leq x < \infty)$, the probability distribution is such that

$$P(X \leq x) = 1 - e^{-\beta x^2}, \ \textit{where } \beta > 0$$

Find the median of the distribution. Also if $m$, $m_0$ and $\sigma$ denote the mean, mode and standard deviation respectively of the distribution, prove that

$$2m_0^2 - m^2 = \sigma^2 \text{ and } m_0 = m\sqrt{2/\pi}$$

What is the sign of skewness of the distribution ?

8. *(a)* Two dice are rolled, $S = \{(a, b) \mid a, b = 1, 2 \ldots, 6\}$. Let $X$ denote the sum of the two faces and $Y$ the absolute value of their difference, *i.e.*, $X$ is distributed over the integers 2, 3, ...., 12 and $Y$ over 0, 1, 2, ..., 5. Assuming the dice are fair, find the probabilities that *(i)* $X = 5 \cap Y = 1$, *(ii)* $X = 7 \cap Y \geq 3$, *(iii)* $X = Y$, and *(iv)* $X + Y = 4 \cap X - Y = 2$.

Ans. *(i)* $1/18$, *(ii)* $1/9$, *(iii)* 0 and *(iv)* $1/18$.

9. The joint probability density function of the two-dimensional variable $(X, Y)$ is of the form

$$f(x, y) = k e^{-(x+y)}, \ 0 \leq y < x < \infty$$
$$= 0, \ \textit{elsewhere}$$

*(i)* Determine the constant $k$. *(ii)* Find the conditional probability density function $f_1(x \mid y)$ and *(iii)* Compute $P(Y \geq 3)$.

**[ Sardar Patel Univ. B.Sc., 1986 ]**

*(iv)* Find the marginal frequency function $f_1(x)$ of $X$.

*(v)* Find the marginal frequency function $f_2(y)$ of $Y$.

*(vi)* Examine if $X, Y$ are independent.

*(vii)* Find the conditional frequency function of $Y$ given $X = 2$.

**Ans.** (i) $k = 1$, (ii) $f_1(x \mid y) = e^{-x}$, (iii) $e^{-3}$.

**10.** Let

$$f(x, y) = \begin{cases} \dbinom{y}{x} p^x (1-p)^{y-x} \dfrac{e^{-\lambda} \lambda^y}{y!} \\ \qquad ; x = 0, 1, 2, \ldots; \ y = 0, 1, 2, \ldots; \ with \ y \geq x \\ 0, \ elsewhere \end{cases}$$

Find the marginal density function of $X$ and the marginal density function of $Y$. Also determine whether the random variables $X$ and $Y$ are independent.

**[I.S.I., 1987]**

**11.** Consider the following function :

$$f(x \mid y) = \begin{cases} \dfrac{y^x e^{-y}}{x!}, & x = 0, 1, 2, \ldots \\ 0, & otherwise \end{cases}$$

*(i)* Show that $f(x \mid y)$ is the conditional probability function of $X$ given $Y$; $y \geq 0$.

*(ii)* If the marginal p.d.f. of $Y$ is

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda x}, & x > 0. \\ 0 & x \leq 0, \ \lambda > 0 \end{cases}$$

what is the joint p.d.f. of $X$ and $Y$ ?

*(iii)* Obtain the marginal probability function of $X$.

**[Delhi Univ. M.A.(Econ.), 1989]**

**12.** The probability density function of $(x_1, x_2)$ is given as

$$f(x_1, x_2) = \begin{cases} \theta_1 \theta_2 e^{-\theta_1 x_1 - \theta_2 x_2} & \text{if } x_1, x_2 > 0 \\ 0 & otherwise. \end{cases}$$

Find the density function of $(y_1, y_2)$ where

$$y_1 = \frac{2x_1}{x_2} + 1, \quad y_2 = 3x_1 + x_2 \ \text{almost everywhere.}$$

**[Punjab Univ. M.A.(Econ.), 1992]**

**13.** *(a)* Let $X_1, X_2$ be a random sample of size 2 from a distribution with probability density function,

$$f(x) = e^{-x}, \ 0 < x < \infty$$

$= 0$, *elsewhere*

Show

$$Y_1 = X_1 + X_2 \text{ and } Y_2 = \frac{X_1}{X_1 + X_2}$$

are independent. [Sardar Patel Univ. B.Sc., Sept. 1986]

*(b)* $X_1$, $X_2$, $X_3$ denote random sample of size 3 drawn from the distribution:

$$f(x) = e^{-x},\ 0 < x < \infty$$
$$= 0,\ \textit{elsewhere}$$

Show that

$$Y_1 = \frac{X_1}{X_1 + X_2},\ Y_2 = \frac{X_1 + X_2}{X_1 + X_2 + X_3} \text{ and } Y_3 = X_1 + X_2 + X_3$$

are mutually independent.

14. If the probability density function of the random varaibles $X$ and $Y \mid X$ is given by

$$f(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & \textit{elsewhere} \end{cases}$$

and $$f_{Y|X}(y \mid x) = \begin{cases} \dfrac{e^{-x} x^y}{y!}, & y \geq 0 \\ 0, & \textit{elsewhere} \end{cases}$$

respectively, find the probability density function of the random variable Y.

[Jiwaji Univ. M.Sc., 1987]

15. *(a)* The random variable $X$ and $Y$ have a joint p.d.f. $f(x, y)$ given by

$$f(x, y) = g(x + y),\quad x > 0,\ y > 0$$
$$= 0,\quad \text{otherwise.}$$

Obtain the distribution function $H(z)$ of $Z = X + Y$ and hence show that its p.d.f. is

$$h(z) = z\,g(z),\quad z > 0$$
$$= 0\quad z \leq 0.$$

*(b)* The joint density function of two random variables is given by

$f(x, y) = e^{-(x+y)}$ ; $x > 0$, $y > 0$. Show that the p.d.f. of

$$U = \frac{X + Y}{2} \text{ is } g(u) = 4\,u\,e^{-2u}$$

[Calicut Univ. B.Sc., 1986]

16. The time $X$ taken by a garage to repair a car is a continuous random variable with probability density function

$$f_1(x) = \begin{cases} \frac{3}{4}x(2-x), & 0 \le x \le 2 \\ 0, & elsewhere \end{cases}$$

If, on leaving his car, a motorist goes to keep an engagement, lasting for a time $Y$, where $Y$ is a continuous random variable, independent of $X$, with probability function

$$f_2(y) = \begin{cases} \frac{1}{2}y; & 0 \le y \le 2 \\ 0, & elsewhere; \end{cases}$$

determine the probability that the car will not be ready on his return.

**[Calcutta Univ. B.A.(Hons.), 1988]**

17. If $X$ and $Y$ are two independent random variables such that

$$f(x) = e^{-x}, \ x \ge 0 \quad \text{and} \quad g(y) = 3e^{-3y}, \ y \ge 0;$$

find the probability distribution of $Z = X/Y$.

**[Madurai Univ. B.Sc., Oct. 1987]**

18. The random variables $X$ and $Y$ are independent and their probability density functions are, respectively given by

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{\sqrt{1-x^2}}, \ |x| < 1 \quad \text{and} \quad g(y) = y e^{-y^2/2}, \ y > 0.$$

Find the joint probability density of $Z$ and $W$ where $Z = XY$ and $W = X$. Deduce the probability density of $Z$. **[Calcutta Univ. B.Sc.(Hons.), 1985]**

CHAPTER SIX

# Mathematical Expectation, Generating Functions and Law of Large Numbers

**6·1. Mathematical Expectation.** Let $X$ be a random variable (r.v.) with p.d.f. (p.m.f.) $f(x)$. Then its mathematical expectation, denoted by $E(X)$ is given by :

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx, \quad (\text{ for continuous } r.v.) \qquad \ldots(6\cdot1)$$

$$= \sum_x x\, f(x), \quad (\text{ for discrete } r.v.) \qquad \ldots(6\cdot1a)$$

provided the righthand integral or series is absolutely convergent, *i.e.*, provided

$$\int_{-\infty}^{\infty} |x f(x)|\, dx = \int_{-\infty}^{\infty} |x|\, f(x)\, dx < \infty \qquad \ldots(6\cdot2)$$

or
$$\sum_x |x\, f(x)| = \sum_x |x|\, f(x) < \infty \qquad \ldots(6\cdot2a)$$

**Remarks. 1.** Since absolute convergence implies ordinary convergence, if (6·2) or (6·2*a*) holds then the integral or series in (6·1) and (6·1*a*) also exists, *i.e.*, has a finite value and in that case we define $E(X)$ by (6·1) or (6·1*a*). It should be clearly understood that although $X$ has an expectation only if L.H.S. in (6·2) or (6·2*a*) exists, *i.e.*, converges to a finite limit, its value is given by (6·1) or (6·1*a*).

**2.** $E(X)$ exists iff $E|X|$ exists.

**3.** The expectation of a random variable is thought of as a long-term average. [See Remark to Example (6·2*a*), **page 6·19.**].

**4. Expected value and variance of an Indicator Variable.** Consider the indicator variable : $X = I_A$ so that

$$X = 1 \quad \text{if } A \textit{ happens}$$
$$= 0 \quad \text{if } \bar{A} \textit{ happens}$$

$$\therefore \quad E(X) = 1 \;\; P(X=1) + 0 \,.\, P(X=0)$$
$$\Rightarrow \quad E(I_A) = 1 \;\; P[I_A = 1] + 0 \,.\, P[I_A = 0]$$
$$\Rightarrow \quad E(I_A) = P(A)$$

This gives us a very useful tool to find $P(A)$, rather than to evaluate $E(X)$.

Thus
$$P(A) = E(I_A) \qquad \ldots(6\cdot2b)$$

For illustration of this result, see Example 6·14, page 6·27.

$$E(X^2) = 1^2 \,.\, P(X=1) + 0^2 \,.\, P(X=0) = P(I_A = 1) = P(A)$$

$$\therefore \quad Var\, X = E(X^2) - [E(X)]^2 = P(A) - [P(A)]^2$$
$$= P(A)\,[1 - P(A)]$$

$$= P(A)\, P(\bar{A}) \qquad \dots(6\cdot 2c)$$

**Illustrations.** If the r.v. $X$ takes the values $0\,!, 1\,!, 2\,!, \dots$ with probability law

$$P(X = x\,!) = \frac{e^{-1}}{x\,!}\,; \quad x = 0, 1, 2, \dots$$

then
$$\sum_{x=0}^{\infty} x\,!\; P(X = x\,!) = e^{-1} \sum_{x=0}^{\infty} 1$$

which is a divergent series. In this case $E(X)$ does not exist.

More rigorously, let us consider a random variable $X$ which takes the values

$$x_i = (-1)^{i+1}\,(i+1)\,; \quad i = 1, 2, 3, \dots$$

with the probability law

$$p_i = P(X = x_i) = \frac{1}{i\,(i+1)}\,; \quad i = 1, 2, 3, \dots$$

Here
$$\sum_{i=1}^{\infty} x_i\, P(X = x_i) = \sum_{i=1}^{\infty} (-1)^{i+1} \left(\frac{1}{i}\right) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

Using Leibnitz test for alternating series the series on right hand side is conditionally convergent since the terms alternate in sign, are monotonically decreasing and converge to zero. By conditional convergence we mean that although $\sum\limits_{i=1}^{\infty} p_i x_i$ converges, $\sum\limits_{i=1}^{\infty} |p_i\, x_i|$ does not converge. So, rigorously speaking, in the above example $E(X)$ does not exist, although $\sum\limits_{i=1}^{\infty} p_i x_i$ is finite, *viz.*, $\log_e 2$.

As another example, let us consider the r.v. $X$ which takes the values

$$x_k = \frac{(-1)^k \cdot 2^k}{k}\,; \quad k = 1, 2, 3, \dots$$

with probabilities $\quad p_k = 2^{-k}$.

Here also we get

$$\sum_{k=1}^{\infty} x_k\, p_k = \sum_{k=1}^{\infty} \frac{(-1)^k}{k}$$
$$= -\left[1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots\right] = -\log_e 2$$

and
$$\sum_{k=1}^{\infty} |x_k|\, p_k = \sum_{k=1}^{\infty} \frac{1}{k},$$

which is a divergent series. Hence in this case also expectation does not exist.

As an illustration of a continuous r.v. let us consider the r.v. $X$ with p.d.f.

$$f(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2}\ ;\quad -\infty<x<\infty$$

which is p.d.f. of Standard Cauchy distribution. [c.f.s. 8·9].

$$\int_{-\infty}^{\infty}|x|\,f(x)\,dx=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{|x|}{1+x^2}\,dx=\frac{2}{\pi}\int_{0}^{\infty}\frac{x}{1+x^2}\,dx$$

( $\because$ *Integrand is an even function of x* )

$$=\frac{1}{\pi}\left|\log\left(1+x^2\right)\right|_0^{\infty}\to\infty$$

Since this integral does not converge to a finite limit, $E(X)$ does not exist.

**6·2. Expectation of a Function of a Random Variable.** Consider a r.v. $X$ with p.d.f. (p.m.f.) $f(x)$ and distribution function $F(x)$. If $g(.)$ is a function such that $g(X)$ is a r.v. and $E[g(X)]$ exists (*i.e.*, is defined), then

$$E[g(X)]=\int_{-\infty}^{\infty}g(x)\,dF(x)=\int_{-\infty}^{\infty}g(x)f(x)\,dx \qquad ...(6\cdot3)$$

(*For continuous r.v.*)

$$=\sum_x g(x)f(x) \qquad ...(6\cdot3a)$$

(*For discrete r.v.*)

By definition, the expectation of $Y=g(X)$ is

$$E[g(X)]=E(Y)=\int y\,.\,dH_Y(y)=\int y\,h(y)\,dy \qquad ...(6\cdot4)$$

or
$$E(Y)=\sum_y y\,h(y) \qquad ...(6\cdot4a)$$

where $H_Y(y)$ is the distribution function of $Y$ and $h(y)$ is p.d.f. of $Y$.

*[The proof of equivalence of (6·3) and (6·4) is beyond the scope of the book.]*

This result extends into higher dimensions. If $X$ and $Y$ have a joint p.d.f. $f(x,y)$ and $Z=h(x,y)$ is a random variable for some function $h$ and if $E(Z)$ exists, then

$$E(Z)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h(x,y)f(x,y)\,dx\,dy \qquad ...(6\cdot5)$$

or
$$E(Z)=\sum_x\sum_y h(x,y)\,f(x,y) \qquad ...(6\cdot5a)$$

**Particular Cases. 1.** If we take $g(X)=X^r$, $r$ being a positive integer, in (6·3) we get :

$$E(X^r)=\int_{-\infty}^{\infty}x^r\cdot f(x)\,dx \qquad ...(6\cdot5b)$$

which is defined as $\mu_r'$, the $r$th moment (about origin) of the probability distribution.

Thus $\mu_r'$ (*about origin*) $=E(X^r)$. In particular

$\mu_1'$ (*about origin*) $=E(X)$ and $\mu_2'$ (*about origin*) $=E(X^2)$.

Hence $\quad Mean = \bar{x} = \mu_1' \text{ ( about origin )} = E(X)$ ...(6·6)

and $\quad \mu_2 = \mu_2' - \mu_1'^2 = E(X^2) - \{E(X)\}^2$ ...(6·6a)

2. If $g(X) = [X - E(X)]^r = (X - \bar{x})^r$, then from (6·3) we get :

$$E[X - E(X)]^r = \int_{-\infty}^{\infty} [x - E(X)]^r f(x)\, dx = \int_{-\infty}^{\infty} (x - \bar{x})^r f(x)\, dx \quad ...(6·7)$$

which is $\mu_r$, the $r$th moment about mean.

In particular, if $r = 2$, we get

$$\mu_2 = E[X - E(X)]^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x)\, dx \quad ...(6·8)$$

Formulae (6·6a) and (6·8) give the variance of the probability distribution of a r.v. $X$ in terms of expectation.

3. Taking $g(x) = constant = c$, say in (6·3) we get

$$E(c) = \int_{-\infty}^{\infty} c \cdot f(x)\, dx = c \int_{-\infty}^{\infty} f(x)\, dx = c \quad ...(6·9)$$

$$E(c) = c \quad ...(6·9a)$$

**Remark.** The corresponding results for a discrete r.v. $X$ can be obtained on replacing integration by summation ( $\Sigma$ ) over the given range of the variable $X$ in the formulae (6·5) to (6·9).

In the following sections, we shall establish some more results on Expectation in the form of theorems, for continuous r.v.'s only. The corresponding results for discrete r.v.'s can be obtained similarly on replacing integration by summation ( $\Sigma$ ) over the given range of the variable $X$ and are left as an exercise to the reader.

## 6·3. Addition Theorem of Expectation

**Theorem 6·1.** *If $X$ and $Y$ are random variables then*

$$E(X + Y) = E(X) + E(Y), \quad ...(6·10)$$

*provided all the expectations exist.*

**Proof.** Let $X$ and $Y$ be continuous r.v.'s with joint p.d.f. $f_{X,Y}(x, y)$ and marginal p.d.f 's $f_X(x)$ and $f_Y(y)$ respectively. Then by definition :

$$E(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx \quad ...(6·11)$$

$$E(Y) = \int_{-\infty}^{\infty} y\, f_Y(y)\, dy \quad ...(6·12)$$

$$E(X + Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f_{XY}(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\, f_{XY}(x, y)\, dx\, dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y\, f_{XY}(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} x \left[ \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy \right] dx + \int_{-\infty}^{\infty} y \left[ \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx \right] dy$$

$$= \int_{-\infty}^{\infty} x\, f_X(x)\, dx + \int_{-\infty}^{\infty} y\, f_Y(y)\, dy$$

$$= E(X) + E(Y) \qquad \text{[On using (6·11) and (6·12)]}$$

The result in (6·10) can be extended to $n$ variables as given below.

**Theorem 6·1(a).** *The mathematical expectation of the sum of $n$ random variables is equal to the sum of their expectations, provided all the expectations exist.*

*Symbolically, if $X_1, X_2, ..., X_n$ are random variables then*

$$E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n) \qquad \ldots(6·13)$$

or
$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i), \qquad \ldots(6·13a)$$

*if all the expectations exist.*

**Proof.** Using (6·10), for two r.v.'s $X_1$ and $X_2$ we get :

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

$\Rightarrow$ (6·13) is true for $n = 2$. ...(*)

Let us now suppose that (6·13) is true for $n = r$ (say), so that

$$E\left(\sum_{i=1}^{r} X_i\right) = \sum_{i=1}^{r} E(X_i) \qquad \ldots(6·14)$$

$$E\left(\sum_{i=1}^{r+1} X_i\right) = E\left[\sum_{i=1}^{r} X_i + X_{r+1}\right]$$

$$= E\left(\sum_{i=1}^{r} X_i\right) + E(X_{r+1}) \qquad \text{[Using (6·10)]}$$

$$= \sum_{i=1}^{r} E(X_i) + E(X_{r+1}) \qquad \text{[Using (6·14)]}$$

$$= \sum_{i=1}^{r+1} E(X_i)$$

Hence if (6·13) is true for $n = r$, it is also true for $n = r + 1$. But we have proved in (*) above that (6·13) is true for $n = 2$. Hence it is true for $n = 2 + 1 = 3$ ; $n = 3 + 1 = 4$ ; ... and so on. Hence by the principle of mathematical Introduction (6·13) is true for all positive integral values of $n$.

## 6·4. Multiplication Theorem of Expectation

**Theorem 6·2.** *If $X$ and $Y$ are independent random variables, then*

$$E(XY) = E(X) \cdot E(Y) \qquad ...(6\cdot15)$$

**Proof.** Proceeding as in Theorem 6·1, we have :

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f_{XY}(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f_X(x) f_Y(y)\, dx\, dy$$

*[Since $X$ and $Y$ are independent]*

$$= \int_{-\infty}^{\infty} x\, f_X(x)\, dx \int_{-\infty}^{\infty} y\, f_Y(y)\, dy$$

$$= E(X) E(Y), \qquad \textit{[Using (6·11) and (6·12)]}$$

provided $X$ and $Y$ are independent.

**Generalisation to n-variables.**

**Theorem 6·2(a).** *The mathematical expectation of the product of a number of independent random variables is equal to the product of their expectations. Symbolically, if $X_1, X_2, ..., X_n$ are $n$ independent random variables, then*

$$\left.\begin{aligned} E(X_1 X_2 \dots X_n) &= E(X_1) E(X_2) \dots E(X_n) \\ \text{i.e.,} \quad E\left(\prod_{i=1}^{n} X_i\right) &= \prod_{i=1}^{n} E(X_i) \end{aligned}\right\} \qquad ...(6\cdot16)$$

*provided all the expectations exist.*

**Proof.** Using (6·15), for two *independent* random variables $X_1$ and $X_2$, we get:

$$E(X_1 X_2) = E(X_1) E(X_2).$$

$\Rightarrow$ (6·16) is true for $n = 2$. ...(*)

Let us now suppose that (6·16) is true for $n = r$, (say) so that :

$$E\left(\prod_{i=1}^{r} X_i\right) = \prod_{i=1}^{r} E(X_i) \qquad ...(6\cdot17)$$

$$E\left(\prod_{i=1}^{r+1} X_i\right) = E\left(\prod_{i=1}^{r} X_i \times X_{r+1}\right)$$

$$= E\left(\prod_{i=1}^{r} X_i\right) E(X_{r+1}) \qquad \text{[Using (6·15)]}$$

$$= \prod_{i=1}^{r} (E X_i)\, E(X_{r+1}) \qquad \text{[Using (6·17)]}$$

$$= \prod_{i=1}^{r+1} (E X_i)$$

Hence if (6·16) is true for $n = r$, it is also true for $n = r + 1$. Hence using (*), by the principle of mathematical induction we conclude that (6·16) is true for all positive integral values of $n$.

**Theorem 6·3.** *If $X$ is a random variable and '$a$' is constant, then*

(i) $E[a\Psi(X)] = a\, E[\Psi(X)]$ ...(6·18)

(ii) $E[\Psi(X) + a] = E[\Psi(X)] + a,$ ...(6·19)

*where $\Psi(X)$, a function of $X$, is a r.v. and all the expectations exist.*

**Proof.**

(i) $$E[a\Psi(X)] = \int_{-\infty}^{\infty} a\Psi(x) \cdot f(x)\, dx = a\int_{-\infty}^{\infty} \Psi(x) f(x)\, dx = a\, E[\Psi(X)]$$

(ii) $$E[\Psi(X) + a] = \int_{-\infty}^{\infty} [\Psi(x) + a]\, f(x)\, dx$$

$$= \int_{-\infty}^{\infty} \Psi(x)\, f(x)\, dx + a\int_{-\infty}^{\infty} f(x)\, dx$$

$$= E[\Psi(X)] + a \qquad \left(\because \int_{-\infty}^{\infty} f(x)\, dx = 1\right)$$

**Cor.** (i) If $\Psi(X) = X$, then

$E(aX) = a\,E(X)$ and $E(X + a) = E(X) + a$ ...(6·20)

(ii) If $\Psi(X) = 1$, then $E(a) = a$. ...(6·21)

**Theroem 6·4.** *If $X$ is a random variable and a and b are constants, then*

$$E(aX + b) = a\, E(X) + b \qquad ...(6·22)$$

*provided all the expectations exist.*

**Proof.** By definition, we have

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b)\, f(x)\, dx$$

$$= a\int_{-\infty}^{\infty} x\, f(x)\, dx + b\int_{-\infty}^{\infty} f(x)\, dx$$

$$= a\, E(X) + b$$

**Cor. 1.** If $b = 0$, then we get

$$E(aX) = a \cdot E(X) \qquad ...(6·22a)$$

**Cor. 2.** Taking $a = 1$, $b = -\bar{X} = -E(X)$, we get

$$E(X - \bar{X}) = 0$$

**Remark.** If we write,

$$g(X) = aX + b \qquad ...(6·23)$$

then $$g[E(X)] = a\,E(X) + b \qquad ...(6·23a)$$

Hence from (6·22) and (6·23*a*) we get

$$E[g(X)] = g[E(X)] \qquad \text{....(6·24)}$$

Now (6·23) and (6·24) imply *that expectation of a linear function is the same linear function of the expectation.* The result, however, is not true if $g(\cdot)$ is not linear. For instance

$$E(1/X) \neq (1/EX) \quad ; \quad E(X^{\frac{1}{2}}) \neq [E(X)]^{\frac{1}{2}},$$
$$E[\log(X) \neq \log[E(X)] \quad ; \quad E(X^2) \neq [E(X)]^2,$$

since all the functions stated above are non-linear. As an illustration, let us consider a random varible $X$ which assumes only two values +1 and −1, each with equal probability $\frac{1}{2}$. Then

$$E(X) = 1 \times \tfrac{1}{2} + (-1) \times \tfrac{1}{2} = 0.$$

and $$E(X^2) = 1^2 \times \tfrac{1}{2} + (-1)^2 \times \tfrac{1}{2} = 1.$$

Thus $$E(X^2) \neq [E(X)]^2$$

For a non-linear function $g(X)$, it is difficult to obtain expressions for $E[g(X)]$ in terms of $g[E(X)]$, say, for $E[\log(X)]$ or $E(X^2)$ in terms of $\log[E(X)]$ or $[E(X)]^2$. However, some results in the form of inequalities between $E[g(X)]$ and $g[E(X)]$ are available, as discussed in Theorem 6·12 (Jenson's Inequality) page 6·15.

**6·5. Expectation of a Linear Combination of Random Variables**

Let $X_1, X_2, ..., X_n$ be any $n$ random variables and if $a_1, a_2, ..., a_n$ are any $n$ constants, then

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i) \qquad \text{...(6·25)}$$

provided all the expectations exist.

**Proof.** The result is obvious from (6·13) and (6·20).

**Theorem 6·5 (a).** *If* $X \geq 0$ *then* $E(X) \geq 0$.

**Proof.** If $X$ is a continuous random variable s.t. $X \geq 0$ then

$$E(X) = \int_{-\infty}^{\infty} x \cdot p(x)\, dx = \int_{0}^{\infty} x \cdot p(x)\, dx > 0,$$

$$[\because \text{If } X \geq 0,\ p(x) = 0 \text{ for } x < 0]$$

provided the expectation exists.

**Theorem 6·5 (b).** *Let* $X$ *and* $Y$ *be two random variables such that* $Y \leq X$ *then*

$$E(Y) \leq E(X),$$

*provided the expectations exist.*

**Proof.** Since $Y \leq X$, we have the r.v.

$$Y - X \leq 0 \quad \Rightarrow \quad X - Y \geq 0$$

Hence $$E(X - Y) \geq 0 \quad \Rightarrow \quad E(X) - E(Y) \geq 0$$

$$\Rightarrow \quad E(X) \geq E(Y) \quad \Rightarrow \quad E(Y) \leq E(X),$$

as desired.

**Theorem 6·6.** $|E(X)| \leq E|X|,$ ...(6·26)

*provided the expectations exist.*

**Proof.** Since $X \leq |X|$, we have by Theorem 6·5(*b*)

$$E(X) \leq E|X| \qquad ...(*)$$

Again since $-X \leq |X|$, we have by Theorem 6·5(*b*)

$$E(-X) \leq E|X|$$

$$\Rightarrow \quad -E(X) \leq E|X| \qquad ...(**)$$

From (*) and (**), we get the desired result $|E(X)| \leq E|X|$.

**Theorem 6·7.** *If $\mu_r'$ exists, then $\mu_s'$ exists for all $1 \leq s \leq r$. Mathematically, if $E(X^r)$ exists, then $E(X^s)$ exists for all $1 \leq s \leq r$, i.e.,*

$$E(X^r) < \infty \Rightarrow E(X^s) < \infty \quad \forall \quad 1 \leq s \leq r \qquad ...(6·27)$$

**Proof.** $$\int_{-\infty}^{\infty} |x|^s \, dF(x) = \int_{-1}^{1} |x|^s \, dF(x) + \int_{|x|>1} |x|^s \, dF(x)$$

If $s < r$, then $|x|^s < |x|^r$ for $|x| > 1$.

$$\therefore \int_{-\infty}^{\infty} |x|^s \, dF(x) \leq \int_{-1}^{1} |x|^s \, dF(x) + \int_{|x|>1} |x|^r \, dF(x)$$

$$\leq \int_{-1}^{1} dF(x) + \int_{|x|>1} |x|^r \, dF(x),$$

since for $-1 < x < 1$, $|x|^s < 1$.

$$\therefore \quad \int_{-\infty}^{\infty} |x|^s \, dF(x) \leq 1 + E|X|^r < \infty$$

$$\Rightarrow \quad E(x^s) \text{ exists } \forall \ 1 \leq s \leq r \qquad [\because E(X^r) \text{ exists}]$$

**Remark.** The above theorem states that if the moments of a specified order exist, then all the lower order moments automatically exist. However, the converse is not true, *i.e.*, we may have distributions for which all the moments of a specified order exist but no higher order moments exist. For example, for the r.v. with p.d.f.

$$p(x) = 2/x^3 \ ; \quad x \geq 1$$
$$= 0 \ ; \quad x < 1$$

we have :

$$E(X) = \int_1^{\infty} x \, p(x) \, dx = 2\int_1^{\infty} x^{-2} \, dx = \left|\left(\frac{-2}{x}\right)\right|_1^{\infty} = 2$$

$$E(X^2) = \int_1^\infty x^2\, p(x)\, dx = 2\int_1^\infty \frac{1}{x}\, dx = \infty$$

Thus for the above distribution, 1st order moment (mean) exists but 2nd order moment (variance) does not exist.

As another illustration, consider a r.v. $X$ with p.d.f.

$$p(x) = \frac{(r+1)\, a^{r+1}}{(x+a)^{r+2}} \;;\; x \geq 0 \;;\quad a > 0$$

$$\mu_r' = E(X^r) = (r+1)\, a^{r+1} \int_0^\infty \frac{x^r}{(x+a)^{r+2}}\, dx$$

Put $x = ay$ and using Beta integral :

$$\int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}} = \beta(m, n),$$

we shall get on simplification :

$$\mu_r' = (r+1)\, a^r \cdot \beta(r+1, 1) = a^r$$

However,

$$\mu_{r+1}' = E\left(X^{r+1}\right) = (r+1)\, a^{r+1} \int_0^\infty \frac{x^{r+1}}{(x+a)^{r+2}}\, dx \to \infty,$$

as the intergal is not convergent. Hence in this case only the moments up to $r$th order exist and higher order moments do not exist.

**Theorem 6·8.** *If $X$ is a random variable, then*

$$V(aX+b) = a^2 V(X), \qquad \ldots(6{\cdot}28)$$

*where a and b are constants.*

**Proof.** Let $Y = aX + b$

Then $E(Y) = a\, E(X) + b$

$\therefore$ $Y - E(Y) = a\{X - E(X)\}$

Squaring and taking expectation of both sides, we get

$$E\{Y - E(Y)\}^2 = a^2 E\{X - E(X)\}^2$$

$\Rightarrow$ $V(Y) = a^2 V(X)$ $\Rightarrow$ $V(aX+b) = a^2 V(X)$,

where $V(X)$ is written for variance of $X$.

**Cor.** *(i)* If $b = 0$, then $V(aX) = a^2 V(X)$ ...(6·28*a*)

$\Rightarrow$ Variance is not independent of change of scale .

*(ii)* If $a = 0$, then $V(b) = 0$ ...(6·28*b*)

$\Rightarrow$ Variance of a constant in zero.

*(iii)* If $a = 1$, then $V(X+b) = V(X)$ ...(6·28*c*)

$\Rightarrow$ Variance is independent of change of origin.

**6·6. Covariance:** *If X and Y are two random variables, then covariance between them is defined as*

$$Cov\ (X, Y) = E\left[\{X - E(X)\}\{Y - E(Y)\}\right] \qquad ...(6\cdot29)$$
$$= E\left[XY - X E(Y) - Y E(X) + E(X) E(Y)\right]$$
$$= E(XY) - E(Y) E(X) - E(X) E(Y) + E(X) E(Y)$$
$$= E(XY) - E(X) E(Y) \qquad ...(6\cdot29a)$$

If $X$ and $Y$ are independent then $E(XY) = E(X) E(Y)$ and hence in this case

$$Cov\ (X, Y) = E(X) E(Y) - E(X) E(Y) = 0 \qquad ...(6\cdot29b)$$

**Remarks. 1.** $Cov\ (aX, bY) = E\left[\{aX - E(aX)\}\{bY - E(bY)\}\right]$
$$= E\left[a\{X - E(X)\} b \{Y - E(Y)\}\right]$$
$$= ab\, E\left[\{X - E(X)\}\{Y - E(Y)\}\right]$$
$$= ab\ Cov\ (X, Y) \qquad ...(6\cdot30)$$

**2.** $Cov\ (X + a,\ Y + b) = Cov\ (X, Y) \qquad ...(6\cdot30a)$

**3.** $Cov\left(\dfrac{X - \overline{X}}{\sigma_X}, \dfrac{Y - \overline{Y}}{\sigma_Y}\right) = \dfrac{1}{\sigma_X\ \sigma_Y}\ Cov\ (X, Y) \qquad ...(6\cdot30b)$

**4.** Similarly, we shall get :

$$Cov\ (aX + b,\ cY + d) = ac\ Cov\ (X, Y) \qquad ...(6\cdot30c)$$
$$Cov\ (X + Y, Z) = Cov\ (X, Z) + Cov\ (Y, Z) \qquad ...(6\cdot30d)$$
$$Cov\ (aX + bY,\ cX + dY) = ac\sigma_X^{\ 2} + bd\sigma_Y^{\ 2} + (ad + bc)\ Cov\ (X, Y) \qquad ...(6\cdot30e)$$

If $X$ and $Y$ are independent, $Cov\ (X, Y) = 0$. [c.f. (6·29b)]

However, the converse is not true.

(For details see Theorem 10·2)

**6·6·1. Correlation Coefficient.** The correlation coefficient ( $\rho_{XY}$ ), between the variables $X$ and $Y$ is defined as :

$$\rho_{XY} = Correlation\ Coefficient\ (X, Y) = \frac{Cov\ (X, Y)}{\sigma_X\ \sigma_Y} \qquad ...(6\cdot30f)$$

For detailed discussion on correlation coefficient, see Chapter 10.

**6·7. Variance of a Linear Combination of Random Variables**

**Theorem 6·9.** *Let $X_1, X_2, ..., X_n$ be n random variables then*

$$V\left[\sum_{i=1}^{n} a_i\ X_i\right] = \sum_{i=1}^{n} a_i^{\ 2}\ V(X_i) + 2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i<j}}^{n} a_i a_j\ Cov\ (X_i, X_j) \qquad ...(6\cdot31)$$

**Proof.** Let $U = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n$

$\therefore \qquad E(U) = a_1 E(X_1) + a_2 E(X_2) + \ldots + a_n E(X_n)$

$\therefore \qquad U - E(U) = a_1 [X_1 - E(X_1)] + a_2 [X_2 - E(X_2)] + \ldots + a_n [X_n - E(X_n)]$

Squaring and taking expectation of both sides, we get

$$E[U - E(U)]^2 = a_1^2 E[X_1 - E(X_1)]^2 + a_2^2 E[X_2 - E(X_2)]^2 + \dots + a_n^2 E[X_n - E(X_n)]^2 + 2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i<j}}^{n} a_i a_j E[\{X_i - E(X_i)\}\{X_j - E(X_j)\}]$$

$$\Rightarrow \quad V(U) = a_1^2 V(X_1) + a_2^2 V(X_2) + \dots + a_n^2 V(X_n) + 2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i<j}}^{n} a_i a_j \, Cov\,(X_i, X_j)$$

$$\Rightarrow V\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 V(X_i) + 2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i<j}}^{n} a_i a_j \, Cov\,(X_i, X_j)$$

**Remarks. 1.** If $a_i = 1$ ; $i = 1, 2, \dots, n$ then

$$V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n) + 2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i<j}}^{n} Cov\,(X_i, X_j) \quad \dots(6\cdot31a)$$

**2.** If $X_1, X_2, \dots, X_n$ are independent (pairwise) then $Cov\,(X_i, X_j) = 0$, $(i \neq j)$. Thus from (6·31) and (6·31$a$), we get

$$\left.\begin{aligned} V(a_1X_1 + a_2X_2 + \dots + a_nX_n) &= a_1^2V(X_1) + a_2^2V(X_2) + \dots + a_n^2V(X_n) \\ \text{and} \quad V(X_1 + X_2 + \dots + X_n) &= V(X_1) + V(X_2) + \dots + V(X_n) \end{aligned}\right\} \quad \dots(6\cdot31b)$$

**3.** If $a_1 = 1 = a_2$ and $a_3 = a_4 = \dots = a_n = 0$, then from (6·31), we get

$$V(X_1 + X_2) = V(X_1) + V(X_2) + 2\,Cov\,(X_1, X_2)$$

Again if $a_1 = 1$, $a_2 = -1$ and $a_3 = a_4 = \dots = a_n = 0$, then

$$V(X_1 - X_2) = V(X_1) + V(X_2) - 2\,Cov\,(X_1, X_2)$$

Thus we have

$$V(X_1 \pm X_2) = V(X_1) + V(X_2) \pm 2\,Cov\,(X_1, X_2) \quad \dots(6\cdot31c)$$

If $X_1$ and $X_2$ are independent, then $Cov\,(X_1, X_2) = 0$ and we get

$$V(X_1 \pm X_2) = V(X_1) + V(X_2) \quad \dots(6\cdot31d)$$

**Theorem 6·10.** *If $X$ and $Y$ are independent random variables then*

$$E[h(X) \cdot k(Y)] = E[h(X)]\, E[k(Y)] \quad \dots(6\cdot32)$$

*where $h(\cdot)$ is a function of $X$ alone and $k(\cdot)$ is a function of $Y$ alone, provided expectations on both sides exist.*

**Proof.** Let $f_X(x)$ and $g_Y(y)$ be the marginal p.d.f.'s of $X$ and $Y$ respectively. Since $X$ and $Y$ are independent, their joint p.d.f. $f_{XY}(x, y)$ is given by

$$f_{XY}(x, y) = f_X(x)\, f_Y(y) \quad \dots(*)$$

By definition, for continuous r.v.'s

$$E[h(X)\cdot k(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)\,k(y)\,f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)\,k(y)\,f(x)\,g(y)\,dx\,dy$$

[From (*)]

Since $E[h(X)\;k(Y)]$ exists, the integral on the right hand side is absolutely convergent and hence by Fubini's theorem for integrable functions we can change the order of integration to get

$$E[h(X)\,k(Y)] = \left[\int_{-\infty}^{\infty} h(x)\,f(x)\,dx\right]\left[\int_{-\infty}^{\infty} k(y)\,g(y)\,dy\right]$$

$$= E[h(X)]\,.\,E[k(Y)],$$

as desired.

**Remark.** The result can be proved for discrete random variables $X$ and $Y$ on replacing integration by summation over the given range of $X$ and $Y$.

**Theorem 6·11. Cauchy-Schwartz Inequality.** *If $X$ and $Y$ are random variables taking real values, then*

$$[E(XY)]^2 \le E(X^2)\,.\,E(Y^2) \qquad \text{...(6.33)}$$

**Proof.** Let us consider a real valued function of the real variable $t$, defined by

$$Z(t) = E(X+tY)^2$$

which is always non-negative, since $(X+tY)^2 \ge 0$, for all real $X$, $Y$ and $t$.

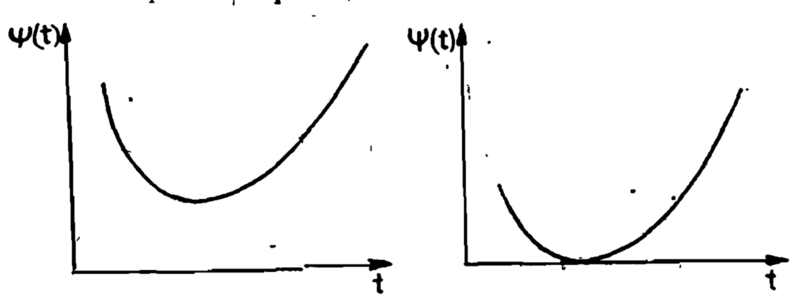
Thus $\quad Z(t) = E(X+tY)^2 \ge 0 \quad \forall \quad t.$

$\Rightarrow \quad Z(t) = E[X^2 + 2t\,XY + t^2Y^2]$

$\qquad = E(X^2) + 2t\,.\,E(XY) + t^2E(Y^2) \ge 0$, for all $t$. $\quad (*)$

Obviously, $Z(t)$ is a quadratic expression in '$t$'.

We know that the quadratic expression of the form :



$\Psi(t) = At^2 + Bt + C \geq 0$ for all $t$, implies that the graph of the function $\Psi(t)$ either touches the $t$-axis at only one point or not at all, as exhibited in the diagrams.

This is equivalent to saying that the discriminant of the function $\Psi(t)$, *viz.*, $B^2 - 4AC \leq 0$, since the condition $B^2 - 4AC > 0$ implies that the function $\Psi(t)$ has two distinct real roots which is a contradiction to the fact that $\Psi(t)$ meets the $t$-axis either at only one point or not at all. Using this result, we get from (*),

$$4 \cdot [E(XY)]^2 - 4 \cdot E(X^2) \cdot E(Y^2) \leq 0$$

$$\Rightarrow \quad [E(XY)]^2 \leq E(X^2) \cdot E(Y^2)$$

**Remarks. 1.** The sign of equality holds in (6·33) or (*) iff

$$E(X + tY)^2 = 0 \quad \forall\ t \quad \Rightarrow \quad P[(X + tY)^2 = 0] = 1,$$

$$\Rightarrow P[X + tY = 0] = 1 \qquad \Rightarrow \quad P\left[Y = -\frac{X}{t}\right] = 1$$

$$\Rightarrow P[Y = AX] = 1 \qquad ; \quad (A = -1/t) \qquad \ldots(6\cdot33b)$$

**2.** If the r.v. $X$ takes the real values $x_1, x_2, \ldots, x_n$ and r.v. $Y$ takes the real values $y_1, y_2, \ldots, y_n$ then Cauchy-Schwartz inequality implies :

$$\left(\frac{1}{n} \sum_{i=1}^{n} x_i\, y_i\right)^2 \leq \left(\frac{1}{n} \sum_{i=1}^{n} x_i^2\right) \cdot \left(\frac{1}{n} \sum_{i=1}^{n} y_i^2\right)$$

$$\Rightarrow \quad \left(\sum_{i=1}^{n} x_i\, y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right) \cdot \left(\sum_{i=1}^{n} y_i^2\right),$$

the sign of equality holding if and only if :

$$\frac{x_i}{y_i} = \textit{constant} = k, \ (\textit{say}) \quad \text{for } \textit{all } i = 1, 2, \ldots, n$$

*i.e.* iff $\dfrac{x_1}{y_1} = \dfrac{x_2}{y_2} = \ldots = \dfrac{x_n}{y_n} = k$, (*say*).

**3.** Replacing $X$ by $|X - E(X)| = |X - \mu_x|$ and taking $Y = 1$ in (6·33) we get

$$[E\,|X - \mu_x|\,]^2 = E\,|X - \mu_x|^2 \cdot E(1)$$

$$\Rightarrow \quad [\textit{ Mean Deviation about mean }]^2 \leq \textit{Variance}\,(X)$$

$$\Rightarrow \quad M.D. \leq S.D. \Rightarrow S.D. \geq M.D. \qquad \ldots(6\cdot33a)$$

## 6·7. Jenson's Inequality

**Continuous Convex Function.** *(Definition).* A continuous function $g(x)$ on the interval $I$ is convex if for every $x_1$ and $x_2$, $(x_1 + x_2)/2 \in I$, we have

$$g\left(\frac{x_1 + x_2}{2}\right) \leq \frac{1}{2}\, g(x_1) + \frac{1}{2}\, g(x_2) \qquad \ldots(6\cdot34)$$

**Remarks. 1.** If $x_1, x_2 \in I$, then $(x_1 + x_2)/2 \in I$.

**2.** Sometimes (6·34) is replaced by the stronger condition :

For $x_1, x_2 \in I$,

$$g[\lambda x_1 + (1 - \lambda) x_2] \leq \lambda g(x_1) + (1 - \lambda) g(x_2) \ ; \ 0 \leq \lambda \leq 1 \qquad \ldots(6\cdot35)$$

(6·34) and (6·35) agree at $\lambda = \frac{1}{2}$.

**3.** If we do not assume the continuity of $g(x)$, then (6·35) is required to define convexity. There are certain *'non-measurable'* non-constant functions $g(\cdot)$ satisfying (6·34) but not (6·35). If $g(\cdot)$ is measurable, then (6·34) and (6·35) are equivalent.

**4.** A function satisfying (6·35) is continuous except possibly at the end points of the interval $I$ (if it has end points).

**5.** If $g$ is twice differentiable, *i.e.*, $g''(x)$ exists for $X \in$ [interior of $I$], and $g''(x) \geq 0$ for such $x$, then $g$ is convex on the interior points.

**6.** For any point $x_0$ interior to $I$, $\exists$ a straight line $y = ax + b$, which passes through $(x_0, g(x_0))$ and satisfies $g(x) \geq ax + b$, for all $x \in I$.

**Theorem 6·12. (Jenson's Inequality).** *If g is continuous and convex function on the interval I, and X is a random variable whose values are in I with probability 1, then*

$$E[g(X)] \geq g[E(X)], \qquad \ldots(6\cdot36)$$

*provided the expectations exist.*

**Proof.** First of all we shall show that $E(X) \in I$.

The various possible cases for $I$ are :

$I = (-\infty, \infty)$; $I = (a, \infty)$; $I = [a, \infty)$; $I = (-\infty, b)$;
$I = (-\infty, b]$, $I = (a, b)$ and variations of this.

If $E(X)$ exists, then $-\infty < E(X) < \infty$.

If $X \geq a$ almost surely *(a.s.), i.e.*, with probability 1, then $E(X) \geq a$.

If $X \leq b$, a.s. then $E(X) \leq b$.

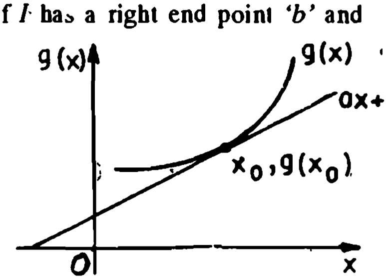
Thus $E(X) \in I$. Now $E(X)$ can be either a left end or a right end point (if end points exist) of I or an interior point of I.

Suppose I has a left end point *'a'*, *i.e.*, $X \geq a$ and $E(X) = a$. Then $X - a \geq 0$ a.s. and $E(X - a) = 0$.

Thus $\quad P(X = a) = 1 \quad$ or $\quad P[(X - a) = 0)] = 1$.

$$\therefore \quad E[g(X)] = E[g(a)] \qquad [\because g(x) = g(a) \text{ a.s.}]$$
$$= g(a) \qquad (\because g(a) \text{ is a constant})$$
$$= g\,E(X)].$$

The result can be established similarly if $I$ has a right end point *'b'* and $E(X) = b$.




Thus we are now required to establish (6·36) when $E(X) = x_0$, is an interior point of I.

Let $ax + b$ pass through the point

$(x_o, g(x_o))$ and let it be below $g$

[c.f. Remark 6 above].

$$\therefore \quad E[g(X)] \geq E(aX+b) = aE(X)+b$$
$$= ax_o + b$$
$$= g(x_o) = g[E(X)]$$
$$\Rightarrow \quad E[g(X)] \geq g[(X)].$$

**Continuous Concave Function.** *(Definition).* A continuous function $g$ is concave on an interval $I$ if $(-g)$ is convex.

**Corollary to Theorem 6·12.** *If $g$ is a continuous and concave function on the interval $I$ and $X$ is a r.v. whose values are in $I$ with probability 1, then*

$$E[g(X)] \leq g[E(X)] \qquad ...(6·37)$$

*provided the expectations exist.*

**Remarks. 1.** Equality holds in Theorem (6·12) and corollary (6·37), if and only if

$$P[g(X) = aX + b] = 1,$$

for some $a$ and $b$.

2. Jenson's inequality extends to random vectors. If $I$ is a convex set in $n$-dimensional Euclidean space, *i.e.*, the interval $I$ in theorem 6·12 is transferred to convex set, $g$ is conotinuous on $I$, (6·34) holds whenever $X_1$ and $X_2$ are any arbitrary vectors in I. The condition $g''(x) \geq 0$ for $x$ interior to $I$ implies

$$\left(\frac{\partial^2 g(x)}{\partial x_i \, \partial x_j}\right) = M(x), \quad \text{(say)},$$

is non-negative definite for all $x$ interior to $I$.

**SOME ILLUSTRATIONS OF JENSON'S INEQUALITY**

1. If $E(X^2)$ exists, then

$$E(X^2) \geq [E(X)]^2, \qquad ...(6·38)$$

since $g(X) = X^2$ is convex function of $X$ as $g''(X) = 2 > 0$.

2. If $X > 0$ a.s. *i.e.*, $X$ assumes only positive values and $E(X)$ and $E(1/X)$ exist then

$$E\left(\frac{1}{X}\right) \geq \frac{1}{E(X)}, \qquad ...(6·38a)$$

because $g(X) = \dfrac{1}{X}$ is a convex function of $X$ since

$$g''(X) = \frac{2}{X^3} > 0, \quad \text{for } X > 0.$$

3. If $X > 0$, a.s. then

$$E(X^{1/2}) \leq [E(X)]^{1/2}, \qquad ...(6·38b)$$

since $g(X) = X^{1/2}$, $X > 0$ is a concave function

as $\quad g''(X) = -\frac{1}{4} X^{-3/2} < 0$, for $X > 0$.

4. If $X > 0$, a.s. then

$$E[\log(X)] \le \log[E(X)], \qquad \text{...(6·38c)}$$

provided the expectations exist, because $\log X$ is a concave function of $X$.

5. Since $g(X) = E(e^{tX})$ is a convex function of $X$ for all $t$ and all $X$, if $E(e^{tX})$ and $E(X)$ exist then

$$E(e^{tX}) \ge e^{tE(X)} \qquad \text{...(6·38d)}$$

If $E(X) = 0$, then

$$M_X(t) = E(e^{tX}) \ge 1, \text{ for all } t.$$

Thus if $M_X(t)$ exists, then it has a lower bound 1, provided $E(X) = 0$. Further, this bound is attained at $t = 0$. Thus $M_X(t)$ has a minimum at $t = 0$.

**6·7·1 ANOTHER USEFUL INEQUALITY.** *Let f and g be monotone functions on some subset of the real line and X be a r.v. whose range is in the subset almost surely (a.s.) If the expectations exist, then*

$$E[f(X)g(X)] \ge E[f(X)]\ E[g(X)] \qquad \text{...(6·39)}$$

or

$$E[f(X)g(X)] \le E[f(X)].E[g(X)] \qquad \text{...(6·39a)}$$

*according as f and g are monotone in the same or in the opposite directions.*

**Proof.** Let us consider the case when both the functions $f$ and $g$ are monotone in the same direction. Let $x$ and $y$ lie in the domain of $f$ and $g$ respectively.

If $f$ and $g$ are both monotonically increasing, then

$$y \ge x \Rightarrow f(y) \ge f(x) \quad \text{and} \quad g(y) \ge g(x)$$

$$\Rightarrow f(y) - f(x) \ge 0 \quad \text{and} \quad g(y) - g(x) \ge 0$$

$$\Rightarrow [f(y) - f(x)].[g(y) - g(x)] \ge 0 \qquad \text{...}(*)$$

If $f$ and $g$ are both monotonically decreasing then for $y \ge x$, we have

$$f(y) \le f(x) \quad \text{and} \quad g(y) \le g(x)$$

$$\Rightarrow f(y) - f(x) \le 0 \quad \text{and} \quad g(y) - g(x) \le 0$$

$$\Rightarrow [f(y) - f(x)].[g(y) - g(x)] \ge 0 \qquad \text{...}(**)$$

Hence if $f$ and $g$ are both monotonic in the same direction, then from (*) and (**), we get the same result, *viz.*,

$$[f(y) - f(x)].[g(y) - g(x)] \ge 0.$$

Let us now consider independently and identically distributed (*i.i.d.*) random variables $X$ and $Y$. Then from above, we get

$$E[(f(Y) - f(X))(g(Y) - g(X))] \ge 0.$$

$$\Rightarrow E[f(Y).g(Y)] - E[f(Y).g(X)] - E[f(X)g(Y)] + E[f(X).g(X)] \ge 0 \qquad \text{...(6·40)}$$

Since $X$ and $Y$ are *i.i.d.* r.v.'s, we have

$$E[f(Y)g(Y)] = E[f(X)g(X)];$$

$$E[f(Y)g(X)] = E[f(Y)]E[g(X)] = E[f(X)]E[g(X)]$$

($\because X$ and $Y$ are independent) ($\because X$ and $Y$ are identical) and $\quad E[f(X)\,g(Y)] = E[f(X)] . E[g(Y)] = E[f(X)] . E[g(X)]$

Substituting in (6·40) we get

$$2\,E[f(X) \cdot g(X)] - 2\,E[f(X)] \cdot E[g(X)] \geq 0\,,$$

$$\Rightarrow \quad E[f(X) \cdot g(X)] \geq E[f(X)] \cdot E[g(X)]$$

which establishes the result in (6·39).

Similarly, (6·39*a*) can be established, if $f$ and $g$ are monotonic in opposite directions, *i.e.*, if $f$ is monotonically increasing (decreasing) and $g$ is monotonically decreasing (increasing). The proof is left as an exercise to the reader.

**SOME ILLUSTRATIONS OF INEQUALITY (6·39).**

1. If $X$ is a r.v. which takes only non-negative values, *i.e.*, if $X \geq 0$ a.s. then for $\alpha > 0$, $\beta > 0$, $f(X) = X^{\alpha}$ and $g(X) = X^{\beta}$ are monotonic in the same direction. Hence if the expectations exist,

$$E(X^{\alpha} \,.\, X^{\beta}) \geq E(X^{\alpha}) . E(X^{\beta})$$

$$\Rightarrow \quad E(X^{\alpha+\beta}) \geq E(X^{\alpha})\, E(X^{\beta})\,;\; \alpha > 0,\; \beta > 0 \qquad \ldots(6\cdot41)$$

In particular, taking $\alpha = \beta = 1$, we get

$$E(X^2) \geq [E(X)]^2\,,$$

a result already obtained in (6·38).

2. If $X \geq 0$, a.s. and $E(X^{\alpha})$ and $E(X^{-1})$ exist, then for $\alpha > 0$, we get from (6·39*a*)

$$E(X^{\alpha} \,.\, X^{-1}) \leq E(X^{\alpha}) . E(X^{-1})$$

$$\Rightarrow \quad E(X^{\alpha})\, E\left(\frac{1}{X}\right) \geq E(X^{\alpha-1})\,;\; \alpha > 0. \qquad \ldots(6\cdot42)$$

In particular with $\alpha = 1$, we get

$$E(X)\, E\left(\frac{1}{X}\right) \geq 1,$$

a result already obtained in (6·38*a*)

Taking $\alpha = 2$ in (6·42), we get

$$E(X^2)\, E\left(\frac{1}{X}\right) \geq E(X)$$

$$\Rightarrow \quad E(X^2) \geq \frac{E(X)}{E\left(\frac{1}{X}\right)} \geq \frac{1}{\left[E\left(\frac{1}{X}\right)\right]^2}\,, \qquad \ldots(6\cdot43)$$

on using (6·38*a*).

Taking $\alpha = 2$ and $\beta = -2$ in $f(X) = X^{\alpha}$, $g(X) = X^{\beta}$ and using (6·39a), we get

$$E(X^2) \,.\, E(X^{-2}) \geq E(X^2 . X^{-2}) = 1.$$

$$\Rightarrow \quad E(X^2) \geq \frac{1}{E(X^{-2})} \qquad \ldots(6\cdot43a)$$

which is a weaker inequality than (6·43).

3. If $M_X(t) = E(e^{tX})$ exists for all $t$ and for some r.v. $X$, then

$$M_{X(u+v)} = E[e^{(u+v)X}] = E(e^{uX} \cdot e^{vX})$$
$$\geq E(e^{uX}) \cdot E(e^{vX})$$
$$= M_X(u) \cdot M_X(v)$$

$\therefore$ $$M_X(u+v) \geq M_X(u) \cdot M_X(v), \quad \text{for } u, v \geq 0.$$

**Example 6·1.** *Let X be a random variable with the following probability distribution :*

| | | | | |
|---|---|---|---|---|
| $x$ | : | $-3$ | $6$ | $9$ |
| $P_r(X = x)$ | : | $1/6$ | $1/2$ | $1/3$ |

Find $E(X)$ and $E(X^2)$ and using the laws of expectation, evaluate $E(2X+1)^2$.

**(Gauhati Univ. B.Sc., 1992)**

**Solution.** $E(X) = \Sigma x \cdot p(x)$

$$= (-3) \times \frac{1}{6} + 6 \times \frac{1}{2} + 9 \times \frac{1}{3} = \frac{11}{2}$$

$$E(X^2) = \Sigma x^2 p(x)$$
$$= 9 \times \frac{1}{6} + 36 \times \frac{1}{2} + 81 \times \frac{1}{3} = \frac{93}{2}$$

$\therefore$ $$E(2X+1)^2 = E[4X^2 + 4X + 1] = 4E(X^2) + 4E(X) + 1$$
$$= 4 \times \frac{93}{2} + 4 \times \frac{11}{2} + 1 = 209$$

**Example 6·2.** (a) *Find the expectation of the number on a die when thrown.*

(b) *Two unbiased dice are thrown. Find the expected values of the sum of numbers of points on them.*

**Solution.** *(a)* Let $X$ be the random variable representing the number on a die when thrown. Then $X$ can take any one of the values 1,2,3,..., 6 each with equal probability 1/6. Hence

$$E(X) = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6} \times 3 + \ldots + \frac{1}{6} \times 6$$
$$= \frac{1}{6}(1 + 2 + 3 + \ldots + 6) = \frac{1}{6} \times \frac{6 \times 7}{2} = \frac{7}{2} \quad \ldots(*)$$

**Remark.** This does not mean that in a random throw of a dice, the player will get the number $(7/2) = 3{\cdot}5$. In fact, one can never get this (fractional) number in a throw of a dice. Rather, this implies that if the player tosses the dice for a "long" period, then on the average toss he will get $(7/2) = 3{\cdot}5$.

*(b)* The probability function of $X$ (the sum of numbers obtained on two dice), is

| Value of X : x | 2 | 3 | 4 | 5 | 6 | 7 | ..... | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|
| Probability | 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | ..... | 2/36 | 1/36 |

$$E(X) = \sum_i p_i x_i$$

$$= 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + 5 \times \frac{4}{36} + 6 \times \frac{5}{36} + 7 \times \frac{6}{36}$$

$$+ 8 \times \frac{5}{36} + 9 \times \frac{4}{36} + 10 \times \frac{3}{36} + 11 \times \frac{2}{36} + 12 \times \frac{1}{36}$$

$$= \frac{1}{36}(2 + 6 + 12 + 20 + 30 + 42 + 40 + 36 + 30 + 22 + 12)$$

$$= \frac{1}{36} \times 252 = 7$$

**Aliter.** Let $X_i$ be the number obtained on the $i$th dice ($i = 1, 2$) when thrown. Then the sum of the number of points on two dice is given by

$$S = X_1 + X_2$$

$$\Rightarrow \quad E(S) = E(X_1) + E(X_2) = \frac{7}{2} + \frac{7}{2} = 7 \qquad \text{[On using (*)]}$$

**Remark.** This result can be generalised to the sum of points obtained in a random throw of $n$ dice. Then

$$E(S) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} (7/2) = \frac{7n}{2}$$

**Example 6·3.** *A box contains $2^n$ tickets among which ${}^nC_i$ tickets bear the number $i$ ; $i = 0, 1, 2, \ldots, n$. A group of $m$ tickets is drawn. What is the expectation of the sum of their numbers ?*

**Solution.** Let $X_i$ ; $i = 1, 2, \ldots, m$ be the variable representing the number on the $i$th ticket drawn. Then the sum '$S$' of the numbers on the tickets drawn is given by

$$S = X_1 + X_2 + \ldots + X_m = \sum_{i=1}^{m} X_i$$

$$\therefore \quad E(S) = \sum_{i=1}^{m} E(X_i)$$

Now $X_i$ is a random variable which can take any one of the possible values 0, 1, 2, ..., $n$ with respective probabilities.

$${}^nC_0/2^n,\ {}^nC_1/2^n,\ {}^nC_2/2^n,\ \ldots,\ {}^nC_n/2^n,$$

$$\therefore \quad E(X_i) = \frac{1}{2^n}\left[1.{}^nC_1 + 2.{}^nC_2 + 3.{}^nC_3 + \ldots + n.{}^nC_n\right]$$

$$= \frac{1}{2^n}\left[1.n + 2.\frac{n(n-1)}{2!} + 3.\frac{n(n-1)(n-2)}{3!} + \ldots + n.1\right]$$

$$= \frac{n}{2^n}\left[1 + (n-1) + \frac{(n-1)(n-2)}{2!} + \ldots + 1\right]$$

$$= \frac{n}{2^n}\left[{}^{n-1}C_0 + {}^{n-1}C_1 + {}^{n-1}C_2 + \ldots + {}^{n-1}C_{n-1}\right]$$

$$= \frac{n}{2^n} \cdot (1+1)^{n-1} = \frac{n}{2}$$

Hence $\quad E(S) = \sum_{i=1}^{m} (n/2) = \dfrac{m \cdot n}{2}$

**Example 6·4.** *In four tosses of a coin, let X be the number of heads. Tabulate the 16 possible outcomes with the corresponding values of X. By simple counting, derive the distribution of X and hence calculate the expected value of X.*

**Solution.** Let $H$ represent a head, $T$ a tail and $X$, the random variable denoting the number of heads.

| *S. No.* | *Outcomes* | *No. of Heads (X)* | *S. No.* | *Outcomes* | *No. of Heads (X)* |
|---|---|---|---|---|---|
| 1 | *H H H H* | 4 | 9 | *H T H T* | 2 |
| 2 | *H H H T* | 3 | 10 | *T H T H* | 2 |
| 3 | *H H T H* | 3 | 11 | *T H H T* | 2 |
| 4 | *H T H H* | 3 | 12 | *H T T T* | 1 |
| 5 | *T H H H* | 3 | 13 | *T H T T* | 1 |
| 6 | *H H T T* | 2 | 14 | *T T H T* | 1 |
| 7 | *H T T H* | 2 | 15 | *T T T H* | 1 |
| 8 | *T T H H* | 2 | 16 | *T T T T* | 0 |

The random variable $X$ takes the values 0, 1, 2, 3 and 4. Since, from the above table, we find that the number of cases favourable to the coming of 0, 1, 2, 3 and 4 heads are 1, 4, 6, 4 and 1 respectively, we have

$$P(X=0) = \frac{1}{16}, \quad P(X=1) = \frac{4}{16} = \frac{1}{4}, \quad P(X=2) = \frac{6}{16} = \frac{3}{8},$$

$$P(X=3) = \frac{4}{16} = \frac{1}{4} \quad \text{and} \quad P(X=4) = \frac{1}{16}.$$

Thus the probability distribution of $X$ can be summarised as follows :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 |
| $p(x)$ : | $\frac{1}{16}$ | $\frac{1}{4}$ | $\frac{3}{8}$ | $\frac{1}{4}$ | $\frac{1}{16}$ |

$$E(X) = \sum_{x=0}^{4} x\,p(x) = 1 \cdot \frac{1}{4} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{16}$$

$$= \frac{1}{4} + \frac{3}{4} + \frac{3}{4} + \frac{1}{4} = 2.$$

**Example 6·5.** *A coin is tossed until a head appears. What is the expectation of the number of tosses required ?* **[Delhi Univ. B.Sc., Oct. 1989]**

**Solution.** Let $X$ denote the number of tosses required to get the first head. Then $X$ can materialise in the following ways :

$$\therefore \qquad E(X) = \sum_{x=1}^{\infty} x\,p(x)$$

| Event | $x$ | Probability $p(x)$ |
|---|---|---|
| $H$ | 1 | $1/2$ |
| $TH$ | 2 | $1/2 \times 1/2 = 1/4$ |
| $TTH$ | 3 | $1/2 \times 1/2 \times 1/2 = 1/8$ |
| $\vdots$ | $\vdots$ | $\vdots$ |

$$= 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 4 \times \frac{1}{16} + \ldots \qquad \ldots(*)$$

This is an arithmetic-geometric series with ratio of GP being $r = 1/2$.

Let $\quad S = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + \ldots$

Then $\quad \frac{1}{2} S = \frac{1}{4} + 2 \cdot \frac{1}{8} + 3 \cdot \frac{1}{16} + \ldots$

$\therefore \quad (1 - \frac{1}{2}) S = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots$

$\Rightarrow \quad \frac{1}{2} S = \dfrac{1/2}{1 - (1/2)} = 1$

[Since the sum of an infinite G.P. with first term $a$ and common ratio $r$ $(< 1)$ is $a/(1 - r)$ ]

$\Rightarrow \quad S = 2$

Hence, substituting in $(*)$, we get

$E(X) = 2$

**Example 6·6.** *What is the expectation of the number of failures preceding the first success in an infinite series of independent trials with constant probability $p$ of success in each trial ?* **[Delhi Univ. B.Sc., Oct. 1991]**

**Solution.** Let the random variable $X$ denote the number of failures preceding the first success. Then $X$ can take the values $0, 1, 2, \ldots, \infty$. We have

$p(x) = P(X = x) = P[x$ failures precede the first success $] = q^x p$

where $q = 1 - p$ is the probability of failure in a trial. Then by def.

$$E(X) = \sum_{x=0}^{\infty} x\, p(x) = \sum_{x=0}^{\infty} x \cdot q^x p = pq \sum_{x=1}^{\infty} x\, q^{x-1}$$

$$= pq [1 + 2q + 3q^2 + 4q^3 + \ldots]$$

Now $1 + 2q + 3q^2 + 4q^3 + \ldots$ is an infinite arithmetic-geometric series.

Let $\quad S = 1 + 2q + 3q^2 + 4q^3 + \ldots$

$\quad qS = q + 2q^2 + 3q^3 + \ldots$

$\therefore \quad (1 - q) S = 1 + q + q^2 + q^3 + \ldots = \dfrac{1}{1 - q}$

$\Rightarrow \quad S = \dfrac{1}{(1 - q)^2}$

$\therefore \qquad 1 + 2q + 3q^2 + 4q^3 + \ldots = \dfrac{1}{(1-q)^2}$

Hence $\qquad E(X) = \dfrac{pq}{(1-q)^2} = \dfrac{pq}{p^2} = \dfrac{q}{p}$

**Example 6·7.** *A box contains 'a' white and 'b' black balls. 'c' balls are drawn. Find the expected value of the number of white balls drawn.*

**[Allahabad Univ. B.Sc., 1989; Indian Forest Service 1987]**

**Solution.** Let a variable $X_i$, associated with $i$th draw, be defined as follows:

$X_i = 1$, if $i$th ball drawn is white

and $\qquad X_i = 0$, if $i$th ball drawn is black

Then the number '$S$' of the white balls among '$c$' balls drawn is given by

$$S = X_1 + X_2 + \ldots + X_c = \sum_{i=1}^{c} X_i \quad \Rightarrow \quad E(S) = \sum_{i=1}^{c} E(X_i)$$

Now $\qquad P(X_i = 1) = P$ (of drawing a white ball) $= \dfrac{a}{a+b}$

and $\qquad P(X_i = 0) = P$ (of drawing a black ball) $= \dfrac{b}{a+b}$

$\therefore \qquad E(X_i) = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = \dfrac{a}{a+b}$

Hence $\qquad E(S) = \displaystyle\sum_{i=1}^{c} \left(\frac{a}{a+b}\right) = \frac{ca}{a+b}$

**Example 6·8.** *Let variate X have the distribution*

$$P(X=0) = P(X=2) = p\,;\; P(X=1) = 1-2p, \quad \text{for } 0 \le p \le \tfrac{1}{2}.$$

*For what p is the Var (X) a maximum ?*

**[Delhi Univ. B.Sc. (Maths Hons.) 1987, 85]**

**Solution.** Here the r.v. $X$ takes the values 0, 1 and 2 with respective probabilities $p$, $1 - 2p$ and $p$, $0 \le p \le \frac{1}{2}$.

$\therefore \qquad E(X) = 0 \times p + 1 \times (1-2p) + 2 \times p = 1$

$\qquad E(X^2) = 0 \times p + 1^2 \times (1-2p) + 2^2 \times p = 1 + 2p$

$\therefore \qquad Var(X) = E(X^2) - [E(X)]^2 = 2p\;;\; 0 \le p \le \frac{1}{2}$

Obviously $Var(X)$ is maximum when $p = \frac{1}{2}$, and

$\qquad [Var(X)]_{max} = 2 \times \frac{1}{2} = 1$

**Example 6·9.** $Var(X) = 0 \Rightarrow P[X = E(X)] = 1$. Comment.

**Solution.** $\quad Var(X) = E[X - E(X)]^2 = 0$

$\Rightarrow \qquad [X - E(X)]^2 = 0$, with probability 1

$\Rightarrow \qquad [X - E(X)] = 0$, with probability 1

$\Rightarrow \qquad P[X = E(X)] = 1$

**Example 6·10.** *Explain by means of an example that a probability distribution is not uniquely determined by its moments.*

**Solution.** Consider a r.v. $X$ with p.d.f. [*c.f.* Log-Normal distribution § 8·2·15

$$f(x) = \frac{1}{\sqrt{2\pi}\,x}\cdot\exp\left[-\tfrac{1}{2}(\log x)^2\right]\ ;\ x>0 \qquad ...(*)$$
$$= 0\ ;\quad \text{otherwise}$$

Consider, another r. v. $Y$ with p.d.f.

$$g(y) = \left[1 + a\sin(2\pi\log y)\right] f(y) = g_a(y)\ ,\ (\text{say}),\quad y>0 \quad ...(**)$$

which, for $-1 \le a \le 1$, represents a family of probability distributions.

$$E(Y^r) = \int_0^\infty y^r\left\{1 + a\sin(2\pi\log y)\right\} f(y)\,dy$$

$$= \int_0^\infty y^r f(y)\,dy + a\,.\int_0^\infty y^r\,.\sin(2\pi\log y)\,f(y)\,dy.$$

$$= EX^r + a\,.\frac{1}{\sqrt{2\pi}}\int_0^\infty y^r\,.\sin(2\pi\log y)\,.\frac{1}{y}\exp\left[-\frac{1}{2}(\log y)^2\right]dy$$

$$= EX^r + \frac{a}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{rz - z^2/2}\ .\sin(2\pi z)\,dz$$

$$\left[\log y = z \ \Rightarrow\ y = e^z\right]$$

$$= EX^r + \frac{a\,.\,e^{r^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-r)^2}\,.\sin(2\pi z)\,dz$$

$$= EX^r + \frac{a\,.\,e^{r^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-y^2/2}\,.\sin(2\pi y)\,dy$$

[$z - r = y \ \Rightarrow\ \sin(2\pi z) = \sin(2\pi r + 2\pi y) = \sin 2\pi y$,
$r$ being a positive integer ].

$$= EX^r,$$

the value of the integral being zero, since the integrand is an odd function of $y$.

$\Rightarrow$ $\quad E(Y^r)$ is independent of 'a' in (**).

Hence, $\{\,g(y) = g_a(y);\ -1 \le a \le 1\,\}$, represents a family of distributions, each different from the other, but having the same moments. This explains that the moments may not determine a distribution uniquely.

**Example 6·11.** *Starting from the origin, unit steps are taken to the right with probability $p$ and to the left with probability $q$ $(= 1 - p)$. Assuming independent movements, find the mean and variance of the distance moved from origin after $n$ steps (Random Walk Problem).*

**Solution.** Let us associate a variable $X_i$ with the $i$th step defined as follows :

$X_i = +1$, if the $i$th step is towards the right,

$= -1$, if the $i$th step is towards the left.

Then $S = X_1 + X_2 + \ldots + X_n = \Sigma X_i$, represents the random distance moved from origin after $n$ steps.

$$E(X_i) = 1 \times p + (-1) \times q = p - q$$

$$E(X_i^2) = 1^2 \times p + (-1)^2 \times q = p + q = 1$$

$$\therefore \quad Var(X_i) = E(X_i^2) - [E(X_i)]^2 = (q+p)^2 - (p-q)^2 = 4pq$$

$$\therefore \quad E(S_n) = \sum_{i=1}^{n} E(X_i) = n(p-q)$$

$$V(S_n) = \sum_{i=1}^{n} V(X_i) = 4npq$$

[ $\because$ *Movements of steps are independent* ].

**Example 6·12.** *Let r.v. $X$ have a density function $f(\cdot)$, cumulative distribution function $F(\cdot)$, mean $\mu$ and variance $\sigma^2$. Define $Y = \alpha + \beta X$, where $\alpha$ and $\beta$ are constants satisfying $-\infty < \alpha < \infty$* and $\beta > 0$.

*(a) Select $\alpha$* and *$\beta$ so that $Y$ has mean 0 and variance 1.*

*(b) What is the correlation coefficient $\rho_{XY}$ between $X$ and $Y$?*

*(c) Find the cumulative distribution function of $Y$ in terms of $\alpha$, $\beta$ and $F(\cdot)$.*

*(d) If $X$ is symmetrically distributed about $\mu$, is $Y$ necessarily symmetrically distributed about its mean?*

**Solution.** *(a)* $E(X) = \mu$, $Var(X) = \sigma^2$. We want $\alpha$ and $\beta$ s.t.

$$E(Y) = E(\alpha + \beta X) = \alpha + \beta\mu = 0 \qquad ...(1)$$

$$Var(Y) = Var(\alpha + \beta X) = \beta^2 . \sigma^2 = 1 \qquad ...(2)$$

Solving *(1)* and *(2)* we get :

$$\beta = 1/\sigma, \ (>0) \text{ and } \alpha = -\mu/\sigma \qquad ...(3)$$

*(b)* $$Cov(X, Y) = E(XY) - E(X)E(Y) = E[X(\alpha + \beta X)]$$

[ $\because E(Y) = 0$ ]

$$= \alpha . E(X) + \beta . E(X^2) = \alpha\mu + \beta[\sigma^2 + \mu^2]$$

$$\therefore \quad \rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{\alpha\mu + \beta[\sigma^2 + \mu^2]}{\sigma . 1} \qquad (\because \sigma_Y = 1)$$

$$= \frac{1}{\sigma^2}\left[-\mu^2 + \sigma^2 + \mu^2\right] = 1 \qquad \text{[On using (3)]}$$

*(c)* Distribution function $G_Y(\cdot)$ of $Y$ is given by :

$$G_Y(y) = P(Y \le y) = P[\alpha + \beta X \le y]$$

$$= P(X \le (y-\alpha)/\beta)$$

$$\Rightarrow \quad G_Y(y) = F_X\left(\frac{y-\alpha}{\beta}\right);$$

*(d)* We have : $$Y = \alpha + \beta X = \frac{1}{\sigma}(X - \mu) = \beta(X - \mu) \qquad \text{[On using (3)]}$$

Since $X$ is given to be symmetrically distributed about mean $\mu$, $(X-\mu)$ and $-(X-\mu)$ have the same distribution.

Hence $Y=\beta(X-\mu)$ and $-Y=-\beta(X-\mu)$ have the same distribution. Since $E(Y)=0$, we conclude that $Y$ is symmetrically distributed about its mean.

**Example 6·13.** *Let $X$ be a r.v. with mean $\mu$ and variance $\sigma^2$. Show that $E(X-b)^2$, as a function of $b$, is minimised when $b=\mu$.*

**Solution.**
$$\begin{aligned} E(X-b)^2 &= E[(X-\mu)+(\mu-b)]^2 \\ &= E(X-\mu)^2+(\mu-b)^2+2(\mu-b)E(X-\mu) \\ &= Var(X)+(\mu-b)^2 \qquad [\because E(X-\mu)=0] \end{aligned}$$

$$\Rightarrow \quad E(X-b)^2 \geq Var(X), \qquad \ldots(*)$$

since $(\mu-b)^2$, being the square of a real quantity is always non-negative.

The sign of equality holds in (*) iff

$$(\mu-b)^2=0 \quad \Rightarrow \quad \mu=b.$$

Hence $E(X-b)^2$ is minimised when $\mu=b$ and its minimum value is $E(X-\mu)^2=\sigma_X{}^2$.

**Remark.** This result states that the sum of squares of deviations is minimum when taken about mean.

[Also see § 2·4, Property 3 of Arithmetic Mean]

**Example 6·14.** *Let $a_1, a_2, \ldots, a_n$ be arbitrary real numbers and $A_1, A_2, \ldots, A_n$ be events. Prove that*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j P(A_i A_j) \geq 0$$

**[Delhi Univ. B.A. (Spl. Course – Stat. Hons.), 1986]**

**Solution.** Let us define the indicator variable :

$$\begin{aligned} X_i = I_{A_i} &= 1 \quad \text{if } A_i \text{ occurs} \\ &= 0 \quad \text{if } \bar{A}_i \text{ occurs}. \end{aligned}$$

Then using (6·2b) :

$$E(X_i) = P(A_i)\,; \quad (i=1,2,\ldots,n) \qquad \ldots(i)$$

Also $\quad X_i X_j = I_{A_i \cap A_j}$

$$\Rightarrow \quad E(X_i X_j) = P(A_i A_j) \qquad \ldots(ii)$$

Consider, for real numbers $a_1, a_2, \ldots, a_n$, the expression $\left(\sum\limits_{i=1}^{n} a_i X_i\right)^2$, which is always non-negative.

$$\Rightarrow \quad \left(\sum_{i=1}^{n} a_i X_i\right)^2 \geq 0$$

$$\Rightarrow \quad \left(\sum_{i=1}^{n} a_i X_i\right)\left(\sum_{j=1}^{n} a_j X_j\right) \geq 0$$

$$\Rightarrow \quad \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j X_i X_j \geq 0, \qquad \ldots(iii)$$

for all $a_i$ 's and $a_j$ 's.

Since expected value of a non-negative quantity is always non-negative, on taking expectations of both sides in *(iii)* and using *(i)* and *(ii)* we get :

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \ E(X_i X_j) \ge 0 \Rightarrow \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \ P(A_i A_j) \ge 0.$$

**Example 6·15.** *In a sequence of Bernoulli trials, let X be the length of the run of either successes or failures starting with the first trial. Find E (X) and V (X).*

**Solution.** Let '$p$' denote the probability of success. Then $q = 1 - p$ is the probability of failuure. $X = 1$ means that we can have any one of the possibilities $SF$ and $FS$ with respective probabilities $pq$ and $qp$.

$$\therefore \quad P(X=1) = P(SF) + P(FS) = pq + qp = 2pq$$

Similarly

$$P(X=2) = P(SSF) + P(FFS) = p^2q + q^2p$$

In general

$$P(X=r) = P[SSS...SF] + P[FFF...FS] = p^r . q + q^r . p$$

$$\therefore \quad E(X) = \sum_{r=1}^{\infty} r P(X=r) = \sum_{r=1}^{\infty} r(p^r . q + q^r . p)$$

$$= pq \ [\sum_{r=1}^{\infty} r . p^{r-1} + \sum_{r=1}^{\infty} r . q^{r-1}]$$

$$= pq \ [(1 + 2p + 3p^2 + ...) + (1 + 2q + 3q^2 + ...)]$$

$$= pq \ [(1-p)^{-2} + (1-q)^{-2}] = pq \ [q^{-2} + p^{-2}]$$

(See Remark to Example 6·17)

$$= pq \left[ \frac{1}{q^2} + \frac{1}{p^2} \right] = \frac{p}{q} + \frac{q}{p}.$$

$$V(X) = E(X^2) - \{E(X)\}^2 = E\{X(X-1)\} + E(X) - \{E(X)\}^2$$

Now

$$E\{X(X-1)\} = \sum_{r=2}^{\infty} r(r-1) P(X=r) = \sum_{r=2}^{\infty} r(r-1)(p^r q + q^r p)$$

$$= \sum_{r=2}^{\infty} r(r-1) p^r q + \sum_{r=2}^{\infty} r(r-1) q^r p$$

$$= p^2 q \sum_{r=2}^{\infty} r(r-1) p^{r-2} + q^2 p \sum_{r=2}^{\infty} r(r-1) q^{r-2}$$

$$= 2p^2 q \sum_{r=2}^{\infty} \frac{r(r-1)}{2} p^{r-2} + 2q^2 p \sum_{r=2}^{\infty} \frac{r(r-1)}{2} q^{r-2}$$

$$= 2p^2 q (1-p)^{-3} + 2q^2 p (1-q)^{-3}$$

$$= 2\left( \frac{p^2}{q^2} + \frac{q^2}{p^2} \right)$$

$$V(X) = 2\left(\frac{p^2}{q^2} + \frac{q^2}{p^2}\right) + \left(\frac{p}{q} + \frac{q}{p}\right) - \left(\frac{p}{q} + \frac{q}{p}\right)^2$$

$$= \left(\frac{p}{q} - \frac{q}{p}\right)^2 + \left(\frac{p}{q} + \frac{q}{p}\right)$$

**Aliter.** Proceed as in Example 6·17.

**Example 6·16.** *A deck of n numbered cards is thoroughly shuffled and the cards are inserted into n numbered cells one by one. If the card number 'i' falls in the cell 'i', we count it as a match, otherwise not. Find the mean and variance of total number of such matches.* **[Delhi Univ. B.Sc., (Stat. Hons.), 1988]**

**Solution.** Let us associate a random variable, $X_i$ with the *i*th draw defined as follows :

$$X_i = \begin{cases} 1, & \text{if the } i\text{th card dealt has the number '}i\text{' on it} \\ 0, & \text{otherwise} \end{cases}$$

Then the total number of matches '$S$' is given by

$$S = X_1 + X_2 + \ldots + X_n = \sum_{i=1}^{n} X_i$$

$$\therefore \quad E(S) = \sum_{i=1}^{n} E(X_i)$$

Now $\quad E(X_i) = 1.P(X_i = 1) + 0.P(X_i = 0) = P(X_i = 1) = \dfrac{1}{n}$

Hence $\quad E(S) = \sum_{i=1}^{n}\left(\dfrac{1}{n}\right) = n \cdot \dfrac{1}{n} = 1$

$$V(S) = V(X_1 + X_2 + \ldots + X_n)$$

$$= \sum_{i=1}^{n} V(X_i) + 2 \sum\sum_{\substack{i,j=1 \\ i<j}}^{n} \text{Cov}(X_i, X_j) \qquad \ldots(1)$$

Now $\quad V(X_i) = E(X_i^2) - \{E(X_i)\}^2$

$$= 1^2 . P(X_i = 1) + 0^2 . P(X_i = 0) - \left(\frac{1}{n}\right)^2$$

$$= \frac{1}{n} - \frac{1}{n^2} = \frac{n-1}{n^2} \qquad \ldots(2)$$

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) \qquad \ldots(3)$$

$$E(X_i X_j) = 1.P(X_i X_j = 1) + 0.P(X_i X_j = 0)$$

$$= \frac{(n-2)!}{n!} = \frac{1}{n(n-1)},$$

since $X_i X_j = 1$ if and only if both card numbers $i$ and $j$ are in their respective matching places and there are $(n-2)!$ arrangements of the remaining cards that correspond to this event.

Substituting in (3), we get

$$\text{Cov}(X_i, X_j) = \frac{1}{n(n-1)} - \frac{1}{n}\cdot\frac{1}{n} = \frac{1}{n^2(n-1)} \qquad ...(4)$$

Substituting from (2) and (4) in (1), we have

$$V(S) = \sum_{i=1}^{n}\left(\frac{n-1}{n^2}\right) + 2\sum_{\substack{i=1\\i<j}}^{n}\sum_{j=1}^{n}\left[\frac{1}{n^2(n-1)}\right]$$

$$= n\left(\frac{n-1}{n^2}\right) + 2.{}^nC_2\,\frac{1}{n^2(n-1)} = \frac{n-1}{n} + \frac{1}{n} = 1$$

**Example 6·17.** *If $t$ is any positive real number, show that the function defined by*

$$p(x) = e^{-t}(1-e^{-t})^{x-1} \qquad ...(*)$$

*can represent a probability function of a random variable $X$ assuming the values 1, 2, 3, ... Find the $E(X)$ and Var $(X)$ of the distribution.*

**[Nagpur Univ. B.Sc., 1988]**

**Solution.** We have

$$e^t > 1,\ \forall\ t > 0 \quad\Rightarrow\quad e^{-t} < 1 \quad\Rightarrow\quad 1 - e^{-t} > 0$$

Also $\quad e^{-t} = \frac{1}{e^t} > 0,\ \forall\ t > 0$

Hence $\quad p(x) = e^{-t}(1-e^{-t})^{x-1} \geq 0\ \ \forall\ \ t > 0, x = 1, 2, 3, \ldots$

Also $\quad \sum_{x=1}^{\infty} p(x) = e^{-t}\sum_{x=1}^{\infty}(1-e^{-t})^{x-1} = e^{-t}\sum_{x=1}^{\infty} a^{x-1}; \qquad [a = 1 - e^{-t}]$

$$= e^{-t}(1 + a + a^2 + a^3 + \ldots) = e^{-t}\times\frac{1}{(1-a)}$$

$$= e^{-t}[1-(1-e^{-t})]^{-1} = e^{-t}(e^{-t})^{-1} = 1$$

Hence $p(x)$ defined in (*) represents the probability function of a r.v. $X$.

$$E(X) = \Sigma x.p(x) = e^{-t}\sum_{x=1}^{\infty} x(1-e^{-t})^{x-1}$$

$$= e^{-t}\sum_{x=1}^{\infty} x.a^{x-1};\ [a = 1-e^{-t}]$$

$$= e^{-t}(1 + 2a + 3a^2 + 4a^3 + \ldots) = e^{-t}(1-a)^{-2} \qquad ...(*)$$

$$= e^{-t}(e^{-t})^{-2} = e^t$$

$$E(X^2) = \Sigma x^2 p(x) = e^{-t}\sum_{x=1}^{\infty} x^2.a^{x-1}$$

$$= e^{-t}[1 + 4a + 9a^2 + 16a^3 + \ldots]$$

$$= e^{-t}(1+a)(1-a)^{-3} = e^{-t}(2-e^{-t})e^{3t}$$

Hence $\quad \text{Var}(X) = E(X^2) - [E(X)]^2 = e^{-t}(2-e^{-t})e^{3t} - e^{2t}$

$$= e^{2t}[(2-e^{-t}) - 1] = e^{2t}(1-e^{-t})$$

$$= e^t(e^t - 1)$$

**Remark.**

(*i*) Consider $\quad S = 1 + 2a + 3a^2 + 4a^3 + \ldots$ (Arithmetico-geometric series)

$$\Rightarrow \quad aS = a + 2a^2 + 3a^3 + \dots$$

$$\Rightarrow \quad (1-a)S = 1 + a + a^2 + a^3 + \dots = 1/(1-a) \Rightarrow S = (1-a)^{-2}$$

$$\sum_{x=1}^{\infty} x\, a^{x-1} = 1 + 2a + 3a^2 + 4a^3 + \dots = (1-a)^{-2} \qquad \dots(*)$$

*(ii)* Consider

$$S = 1 + 2^2 . a + 3^2 . a^2 + 4^2 . a^3 + 5^2 . a^4 + \dots$$

$$\Rightarrow \quad S = 1 + 4a + 9a^2 + 16a^3 + 25a^4 + \dots$$

$$-3a\,S = -3a - 12a^2 - 27a^3 - 48a^4 - \dots$$

$$+3a^2S = +3a^2 + 12a^3 + 27a^4 + \dots$$

$$-a^3S = -a^3 - 4a^4 - \dots$$

Adding the above equations we get :

$$(1-a)^3 S = 1 + a \quad \Rightarrow \quad S = (1+a)(1-a)^{-3} \qquad \dots(**)$$

$$\sum_{x=1}^{\infty} x^2 a^{x-1} = 1 + 4a + 9a^2 + 16a^3 + \dots = (1+a)(1-a)^{-3}$$

The results in (*) and (**) are quite useful for numerical problems and should be committed to memory.

**Example 6·18.** *A man with n keys wants to open his door and tries the keys independently and at random. Find the mean and variance of the number of trials required to open the door (i) if unsuccessful keys are not eliminated from further selection, and (ii) if they are.* **[Rajasthan Univ. B.Sc.(Hons.), 1992]**

**Solution.** *(i)* Suppose the man gets the first success at the $x$th trial, *i.e.*, he is unable to open the door in the first $(x-1)$ trials. If unsuccessful keys are not eliminated then $X$ is a random variable which can take the values 1, 2, 3,.... ad infinity.

Probability of success at the first trial = $1/n$

$\therefore$ Probability of failure at the first trial = $1 - (1/n)$

If unsuccessful keys are not eliminated then the probability of success and consequently of failure is constant for each trial.

Hence $p(x)$ = Probability of 1st success at the $x$th trial

$$= \left(1 - \frac{1}{n}\right)^{x-1} . \frac{1}{n}$$

$$\therefore \quad E(X) = \sum_{x=1}^{\infty} x\, p(x) = \sum_{x=1}^{\infty} x \left(1 - \frac{1}{n}\right)^{x-1} . \frac{1}{n}$$

$$= \frac{1}{n} \sum_{x=1}^{\infty} x A^{x-1}, \text{ where } A = 1 - \frac{1}{n}$$

$$E(X) = \frac{1}{n}[1 + 2A + 3A^2 + 4A^3 + \dots] = \frac{1}{n}(1-A)^{-2}$$

[See (*), Example (6·17)]

$$\frac{1}{n}\left[1-\left(1-\frac{1}{n}\right)\right]^{-2} = n$$

$$E(X^2) = \sum_{x=1}^{\infty} x^2 p(x) = \sum_{x=1}^{\infty} x^2\left(1-\frac{1}{n}\right)^{x-1} \cdot \frac{1}{n}$$

$$= \frac{1}{n} \sum_{x=1}^{\infty} x^2 A^{x-1}$$

$$= \frac{1}{n}[1 + 2^2 . A + 3^2 . A^2 + 4^2 . A^3 + \ldots]$$

$$= \frac{1}{n}(1+A)(1-A)^{-3} \qquad \text{[See (**), Example (6·17)]}$$

$$= \frac{1}{n}\left[1+\left(1-\frac{1}{n}\right)\right]\left[1-\left(1-\frac{1}{n}\right)\right]^{-3}$$

$$= (2n-1)\,n$$

Hence $\quad V(X) = E(X^2) - [E(X)]^2 = (2n-1)\,n - n^2 = n^2 - n = n\,(n-1)$

*(ii)* If unsuccessful keys are eliminated from further selection, then the random variable $X$ will take the values from 1 to $n$. In this case, we have

Probability of success at the first trial = $1/n$

Probability of success at the 2nd trial = $1/(n-1)$

Probability of success at the 3rd trial = $1/(n-2)$

and so on.

Hence probability of 1st succes at 2nd trial = $\left(1-\frac{1}{n}\right)\frac{1}{n-1} = \frac{1}{n}$

Probability of first success at the third trial

$$= \left(1-\frac{1}{n}\right)\left(1-\frac{1}{n-1}\right)\cdot\frac{1}{n-2} = \frac{1}{n},$$

and so on. In general, we have

$$p(x) = \text{Probability of first success at the } x\text{th trial} = \frac{1}{n}$$

$$\therefore \quad E(X) = \sum_{x=1}^{\infty} x\,p(x) = \frac{1}{n}\sum_{x=1}^{n} x = \frac{n+1}{2}$$

$$E(X^2) = \sum_{x=1}^{n} x^2 p(x) = \frac{1}{n}\sum_{x=1}^{n} x^2 = \frac{(n+1)(2n+1)}{6}$$

Hence $\quad V(X) = E(X^2) - [E(X)]^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2$

$$= \frac{n+1}{12}\Big[2(2n+1) - 3(n+1)\Big] = \frac{n^2-1}{12}$$

**Example 6·19.** *In a lottery m tickets are drawn at a time out of n tickets numbered 1 to n. Find the expectation and the variance of the sum S of the numbers on the tickets drawn.* **[Delhi Univ. B.Sc. (Maths Hons.), 1987]**

**Solution.** Let $X_i$ denote the score on the $i$th ticket drawn.

Then $$S = X_1 + X_2 + \ldots + X_m = \sum_{i=1}^{m} X_i,$$

is the total score on the $m$ tickets drawn.

$$\therefore \qquad E(S) = \sum_{i=1}^{m} E(X_i)$$

Now each $X_i$ is a random variable which assumes the values $1, 2, 3, \ldots, n$ each with equal probability $1/n$.

$$\therefore \qquad E(X_i) = \frac{1}{n}(1 + 2 + 3 + \ldots + n) = \frac{(n+1)}{2}$$

Hence $$E(S) = \sum_{i=1}^{m}\left(\frac{n+1}{2}\right) = \frac{m(n+1)}{2}$$

$$V(S) = V(X_1 + X_2 + \ldots + X_m)$$

$$= \sum_{i=1}^{m} V(X_i) + 2\sum_{\substack{i,j\\ i<j}}\sum \text{Cov}(X_i, X_j)$$

$$E(X_i^2) = \frac{1}{n}\left(1^2 + 2^2 + 3^2 + \ldots + n^2\right)$$

$$= \frac{1}{n}\cdot\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

$$\therefore \qquad V(X_i) = E(X_i^2) - [E(X_i)]^2$$

$$= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}$$

Also $$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$$

To find $E(X_i X_j)$ we note that the variables $X_i$ and $X_j$ can take the values as shown below :

| $X_i$ | $X_j$ |
|---|---|
| 1 | $2, 3, \ldots, n$ |
| 2 | $1, 3, \ldots, n$ |
| $\vdots$ | $\vdots$ |
| $n$ | $1, 2, \ldots, (n-1)$ |

Thus the variable $X_i X_j$ can take $n(n-1)$ possible values and $P(X_i = l \cap X_j = k) = \dfrac{1}{n(n-1)}$, $k \neq l$. Hence

$$E(X_i X_j) = \frac{1}{n(n-1)} \begin{bmatrix} 1.2 + 1.3 + \dots\dots + 1.n \\ + 2.1 + 2.3 + \dots\dots + 2.n \\ + \dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots \\ + n.1 + n.2 + \dots + n.(n-1) \end{bmatrix}$$

$$= \frac{1}{n(n-1)} \begin{bmatrix} 1(1+2+3+\dots\dots+n) - 1^2 \\ + 2(1+2+3+\dots\dots+n) - 2^2 \\ + \dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots \\ + n(1+2+\dots\dots \overline{n-1}+n) - n^2 \end{bmatrix}$$

$$= \frac{1}{n(n-1)} \left[ (1+2+3+\dots+n)^2 - (1^2+2^2+\dots+n^2) \right]$$

$$= \frac{1}{n(n-1)} \left[ \left\{ \frac{n(n+1)}{2} \right\}^2 - \frac{n(n+1)(2n+1)}{6} \right]$$

$$= \frac{(n+1)(3n^2-n-2)}{12(n-1)}$$

$$\therefore \qquad \text{Cov}(X_i, X_j) = \frac{(n+1)(3n^2-n-2)}{12(n-1)} - \left( \frac{n+1}{2} \right)^2$$

$$= \frac{(n+1)}{12(n-1)} \left[ 3n^2 - n - 2 - 3(n^2-1) \right]$$

$$= -\frac{(n+1)}{12}$$

Hence $$V(S) = \sum_{i=1}^{m} \left( \frac{n^2-1}{12} \right) + 2 \sum_{i<j=1}^{m}\sum \left\{ -\frac{(n+1)}{12} \right\}$$

$$= \frac{m(n^2-1)}{12} + 2 \cdot \frac{m(m-1)}{2!} \left\{ \frac{-(n+1)}{12} \right\},$$

[since there are ${}^mC_2$ covariance terms in $\text{Cov}(X_i, X_j)$]

$$\therefore \qquad V(S) = \frac{m(n+1)}{12} \Big[ (n-1) - (m-1) \Big] = \frac{m(n+1)(n-m)}{12}$$

**Example 6·20.** *A die is thrown $(n+2)$ times. After each throw a '+' is recorded for 4, 5 or 6 and '–' for 1, 2 or 3, the signs forming an ordered sequence. To each, except the first and the last sign, is attached a characteristic random variable which takes the value 1 if both the neighbouring signs differ from the one between them and 0 otherwise. If $X_1, X_2, \dots, X_n$ are characteristic random variables, find the mean and variance of $X = \sum_{i=1}^{n} X_i$.*

**Solution.** $\quad X = \sum_{i=1}^{n} X_i \quad \Rightarrow \quad E(X) = \sum_{i=1}^{n} E(X_i)$

Now $E(X_i) = 1.P(X_i = 1) + 0.P(X_i = 0) = P(X_i = 1)$

For $X_i = 1$, there are the following two mutually exclusive possibilities :

(i) $-\ +\ -$, (ii) $+\ -\ +$

and since the probability of each sign is $\frac{1}{2}$, we have by addition probability theorem:

$$P(X_i = 1) = P(i) + P(ii) = \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^3 = \frac{1}{4}$$

$$\therefore \quad E(X_i) = \frac{1}{4}$$

Hence $$E(X) = \sum_{i=1}^{n}\left(\frac{1}{4}\right) = \frac{n}{4},$$

$$V(X) = V(X_1 + V_2 + \ldots + X_n)$$

$$= \sum_{i=1}^{n} V(X_i) + 2 \sum\sum_{i<j} \text{Cov}(X_i, X_j) \qquad \ldots(*)$$

Now $$E(X_i^2) = 1^2.P(X_i = 1) + 0^2.P(X_i = 0) = P(X_i = 1) = \frac{1}{4}$$

$$\therefore \quad V(X_i) = E(X_i^2) - [E(X_i)]^2 = \frac{1}{4} - \frac{1}{16} = \frac{3}{16}$$

Now $$E(X_i X_j) = 1.P(X_i = 1 \cap X_j = 1) + 0.P(X_i = 0 \cap X_j = 0)$$

$$4 + 0.P(X_i = 1 \cap X_j = 0) + 0.P(X_i = 0 \cap X_j = 1)$$

$$= P(X_i = 1 \cap X_j = 1)$$

Since there are the following two mutually exclusive possibilities for the event : $(X_i = 1 \cap X_j = 1)$,

(i) $-\ +\ -\ +$

(ii) $+\ -\ +\ -$, we have

$$P(X_i = 1 \cap X_j = 1) = P(i) + P(ii) = \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^4 = \frac{1}{8}$$

$$\therefore \quad \text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$$

$$= \frac{1}{8} - \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$$

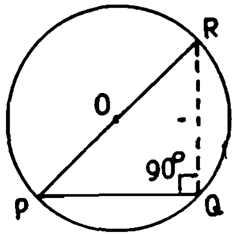
Hence $$V(X) = \sum_{i=1}^{n} (3/16) + 2 \sum\sum_{i<j} \text{Cov}(X_i, X_j) \qquad \text{[From (*)]}$$

$$= \frac{3n}{16} + 2\,[\text{Cov}(X_1, X_2) + \text{Cov}(X_2, X_3) + \ldots + \text{Cov}(X_{n-1}, X_n)]$$

$$= \frac{3n}{16} + 2(n-1) \cdot \frac{1}{16} = \frac{5n-2}{16}$$

**Example 6·21.** *From a point on the circumference of a circle of radius 'a', a chord is drawn in a random direction, (all directions are equally likely). Show that the expected value of the length of the chord is $4a/\pi$ and that the variance of the* *length is* $2a^2\,(1 - 8/\pi^2)$. *Also show that the chance is 1/3 that the length of the chord will exceed the length of the side of an equilateral triangle inscribed in the circle.*

**Solution.** Let $P$ be any point on the circumference of a circle of radius '$a$' and centre '$O$'. Let $PQ$ be any chord drawn at random and let $\angle OPQ = \theta$. Obviously, $\theta$ ranges from $-\pi/2$ to $\pi/2$. Since all the directions are equally likely, the probability differential of $\theta$ is given by the rectangular distribution (*c.f.* Chapter 8):




$$dF(\theta) = f(\theta)\,d(\theta) = \frac{d\theta}{\pi/2 - (-\pi/2)}$$

$$= \frac{d\theta}{\pi}, \quad \frac{-\pi}{2} \le \theta \le \frac{\pi}{2}$$

Now, since $\angle PQR$ is a right angle, (angle in a semi-circle), we have

$$\frac{PQ}{PR} = \cos\theta \;\Rightarrow\; PQ = PR\cos\theta = 2a\cos\theta$$

$$E(PQ) = \int_{-\pi/2}^{\pi/2} (PQ)\,f(\theta)\,d(\theta) = \frac{2a}{\pi}\int_{-\pi/2}^{\pi/2} \cos\theta\,d\theta$$

$$= \frac{2a}{\pi}\Big|\sin\theta\Big|_{-\pi/2}^{\pi/2} = \frac{4a}{\pi}$$

$$E[(PQ)^2] = \int_{-\pi/2}^{\pi/2} [PQ]^2 f(\theta)\,d\theta = \frac{4a^2}{\pi}\int_{-\pi/2}^{\pi/2} \cos^2\theta\,d\theta$$

$$= \frac{4a^2}{\pi}\int_0^{\pi/2} 2\cos^2\theta\,d\theta = \frac{4a^2}{\pi}\int_0^{\pi/2} (1+\cos 2\theta)\,d\theta$$

(Since $\cos^2\theta$ is an even function of $\theta$)

$$= \frac{4a^2}{\pi}\Big|\theta + \frac{\sin 2\theta}{2}\Big|_0^{\pi/2} = \frac{4a^2}{\pi}\cdot\frac{\pi}{2} = 2a^2$$

$$\therefore \quad V(PQ) = E[(PQ)^2] - [E(PQ)]^2 = 2a^2 - \frac{16a^2}{\pi^2} = 2a^2\left(1 - \frac{8}{\pi^2}\right)$$

We know that the length of the side of an equilateral triangle inscribed in a circle of radius '$a$' is $a\sqrt{3}$. Hence

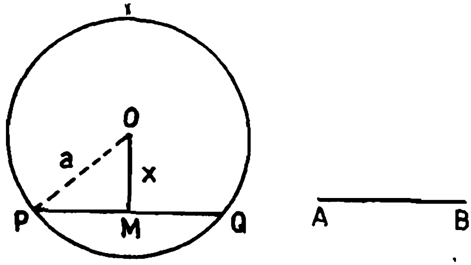
$$P(PQ > a\sqrt{3}) = P(2a\cos\theta > a\sqrt{3}) = P\left(\cos\theta > \frac{\sqrt{3}}{2}\right)$$

$$= P\left(|\theta| < \frac{\pi}{6}\right) = P\left(\frac{-\pi}{6} < \theta < \frac{\pi}{6}\right)$$

$$= \int_{-\pi/6}^{\pi/6} f(\theta)\,d\theta = \frac{1}{\pi}\int_{-\pi/6}^{\pi/6} 1\,.\,d\theta = \frac{1}{\pi}\cdot\frac{\pi}{3} = \frac{1}{3}$$

**Example 6·22.** *A chord of a circle of radius '$a$' is drawn parallel to a given straight line, all distances from the centre of the circle being equally likely. Show that the expected value of the length of the chord is $\pi a/2$ and that the variance of the length is $a^2\,(32 - 3\pi^2)/12$. Also show that the chance is 1/2 that the length of* *the chord will exceed the length of the side of an equilateral triangle inscribed in the circle.*

**Solution.** Let $PQ$ be the chord of a circle with centre $O$ and radius '$a$' drawn at random parallel to the given straight line $AB$. Draw $OM \perp PQ$. Let $OM = x$. Obviously $x$ ranges from $-a$ to $a$. Since all distances from the centre are equally




likely, the probability that a random value of $x$ will lie in the small interval '$dx$' is given by the rectangular distribution [c.f. Chapter 8]:

$$dF(x) = f(x)\,dx = \frac{dx}{a-(-a)} = \frac{dx}{2a}, \quad -a \le x \le a$$

Length of the chord is

$$PQ = 2PM = 2\sqrt{a^2 - x^2}$$

Hence $$E(PQ) = \int_{-a}^{a} PQ\, dF(x) = \frac{2}{2a}\int_{-a}^{a} \sqrt{a^2 - x^2}\,dx$$

$$= \frac{2}{a}\int_0^a \sqrt{a^2 - x^2}\,dx,$$

(since integrand is an even function of $x$).

$$= \frac{2}{a}\left| \frac{1}{2}x\sqrt{a^2 - x^2} + \frac{1}{2}a^2 \sin^{-1}\left(\frac{x}{a}\right)\right|_0^a$$

$$= \frac{2}{a}\cdot\frac{a^2}{2}\cdot\frac{\pi}{2} = \frac{\pi a}{2}$$

$$E[(PQ)^2] = \int_{-a}^{a} (PQ)^2\, dF(x) = \frac{4}{2a}\int_{-a}^{a} (a^2 - x^2)\,dx$$

$$= \frac{4}{a}\int_0^a (a^2 - x^2)\,dx = \frac{4}{a}\left| a^2 x - \frac{x^3}{3}\right|_0^a$$

$$= \frac{4}{a}\cdot\frac{2a^3}{3} = \frac{8a^2}{3}$$

Hence

$$\text{Var (length of chord)} = E[(PQ)^2] - [E(PQ)]^2 = \frac{8a^2}{3} - \frac{\pi^2 a^2}{4}$$

$$= \frac{a^2}{12}(32 - 3\pi^2)$$

The length of the chord is greater than the side of the equilateral triangle inscribed in the circle if

$$2\sqrt{a^2 - x^2} > a\sqrt{3} \quad \Rightarrow \quad 4(a^2 - x^2) > 3a^2$$

*i.e.*, $$x^2 < a^2/4 \quad \Rightarrow \quad |x| < a/2$$

Hence the required probability is

$$P(|x| < a/2) = P\left(-\frac{a}{2} < X < \frac{a}{2}\right) = \int_{-a/2}^{a/2} dF(x)$$

$$= \frac{1}{2a}\int_{-a/2}^{a/2} 1.dx = \frac{1}{2}$$

**Example 6·23.** *Let $X_1, X_2, ..., X_n$ be a sequence of mutually independent random variables with common distribution. Suppose $X_k$ assumes only positive integral values and $E(X_k) = a$, exists; $k = 1, 2, ..., n$. Let $S_n = X_1 + X_2 + ... + X_n$.*

*(i) Show that* $E\left(\frac{S_m}{S_n}\right) = \frac{m}{n}$, *for* $1 \le m \le n$

*(ii) Show that* $E\left(S_n^{-1}\right)$ *exists and*

$$E\left(\frac{S_m}{S_n}\right) = 1 + (m - n)\, a\, E(S_n^{-1}), \text{ for } 1 \le n \le m$$

*(iii) Verify and use the inequality* $x + x^{-1} \ge 2, (x > 0)$ *to show that*

$$E\left(\frac{S_m}{S_n}\right) \ge \frac{m}{n} \text{ for } m, n \ge 1$$

[Delhi Univ. M.Sc. (Stat.), 1988]

**Solution.** *(i)* We have

$$E\left[\frac{X_1 + X_2 + ... + X_n}{X_1 + X_2 + ... + X_n}\right] = E(1) = 1$$

$$\Rightarrow \quad E\left[\frac{X_1 + X_2 + ... + X_n}{S_n}\right] = 1$$

$$\Rightarrow \quad E\left(\frac{X_1}{S_n}\right) + E\left(\frac{X_2}{S_n}\right) + ... + E\left(\frac{X_n}{S_n}\right) = 1$$

Since $X_i$'s, $(i = 1, 2, ..., n)$ are identically distributed random variables, $(X_i/S_n)$, $(i = 1, 2, ..., n)$ are also identically distributed random variables.

$$\therefore \quad n\,E\left(\frac{X_i}{S_n}\right) = 1$$

$$\Rightarrow \quad E\left(\frac{X_i}{S_n}\right) = \frac{1}{n}\,; \quad i = 1, 2, ..., n \qquad ...(*)$$

Now

$$E\left(\frac{S_m}{S_n}\right) = E\left(\frac{X_1 + X_2 + ... + X_m}{S_n}\right) = E\left[\frac{X_1}{S_n} + \frac{X_2}{S_n} + ... + \frac{X_m}{S_n}\right]$$

$$= E\left(\frac{X_1}{S_n}\right) + E\left(\frac{X_2}{S_n}\right) + \dots + E\left(\frac{X_m}{S_n}\right)$$

$$= \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n} \text{ [(}m\text{ times)]} \qquad [Using\ (*)]$$

$$= \frac{m}{n}, (m < n)$$

*(ii)* Since $X_i$'s assume only positive integral values, we have

$$n \le X_1 + X_2 + \dots + X_n < \infty$$

$$\Rightarrow \quad \frac{1}{n} \ge \frac{1}{S_n} > 0 \quad \Rightarrow \quad 0 < S_n^{-1} \le \frac{1}{n}$$

Since $S_n^{-1}$ lies between two finite quantities 0 and $\frac{1}{n}$, we get

$$0 < E(S_n^{-1}) \le \frac{1}{n}$$

Hence $E(S_n^{-1})$ exists.

$$E\left(\frac{S_m}{S_n}\right) = E\left[\frac{X_1 + X_2 + \dots + X_n + X_{n+1} + \dots + X_m}{S_n}\right], \ m \ge n$$

$$= E\left[1 + \frac{X_{n+1}}{S_n} + \dots + \frac{X_m}{S_n}\right]$$

$$= 1 + E\left(\frac{X_{n+1}}{S_n}\right) + \dots + E\left(\frac{X_m}{S_n}\right)$$

Since $X_{n+1}, X_{n+2}, \dots, X_m$ are independent of $S_n = X_1 + X_2 + \dots + X_n$, they are independent of $S_n^{-1}$ also.

$$\therefore \quad E\left(\frac{S_m}{S_n}\right) = 1 + E(X_{n+1}) E(S_n^{-1}) + \dots + E(X_m) E(S_n^{-1})$$

$$= 1 + [E(X_{n+1}) + \dots + E(X_m)] E(S_n^{-1})$$

$$= 1 + (m - n)\, a\, E(S_n^{-1}), \ 1 \le n \le m$$

$$[\because E(X_i) = a \ \forall \ i]$$

*(iii)* Verification of $x + \frac{1}{2} \ge 2, (x > 0)$.

$$x + \frac{1}{2} \ge 2$$

$$\Rightarrow \quad x^2 + 1 \ge 2x \quad \text{(multiplication valid only if } x > 0)$$

$$\Rightarrow \quad (x - 1)^2 \ge 0$$

which is always true for $x > 0$.

If $1 \le m \le n$, result follows from (*).

If $1 \le n \le m$, then using (**), we have to prove that

$$1 + (m-n)\, a\, E(S_n^{-1}) \geq \frac{m}{n}$$

$$\Rightarrow \quad (m-n) a\, E(S_n^{-1}) \geq \frac{m-n}{n}$$

$$\Rightarrow \quad E(S_n^{-1}) \geq \frac{1}{an} \quad \text{...(***)}$$

In (**), taking $x = \dfrac{S_n}{an} > 0$, we get

$$E(x) + E(x^{-1}) \geq 2$$

$$\Rightarrow \quad E\left[\frac{S_n}{an}\right] + E\left[\frac{S_n}{an}\right]^{-1} \geq 2$$

$$\Rightarrow \quad \frac{1}{an} . E(S_n) + an\, E(S_n^{-1}) \geq 2$$

$$\Rightarrow \quad \frac{1}{an} . an + an\, E(S_n^{-1}) \geq 2$$

$$\Rightarrow \quad an\, E(S_n^{-1}) \geq 1$$

$$\Rightarrow \quad E(S_n^{-1}) \geq \frac{1}{an},$$

which was to be proved in (***).

**Example 6·24.** *Let $X$ be a r.v. for which $\beta_1$ and $\beta_2$ exist. Then for any real $k$, prove that :*

$$\beta_2 \geq \beta_1 - (2k + k^2) \quad \text{...(*)}$$

*Deduce that (i) $\beta_2 \geq \beta_1$, (ii) $\beta_2 \geq 1$. When is $\beta_2 = 1$ ?*

**Solution.** Without any loss of generality we can take $E(X) = 0$. [If $E(X) \neq 0$, then we may start with the random variable $Y = X - E(X)$ so that $E(Y) = 0$.]

Consider the real valued function of the real variable $t$ defined by :

$$Z(t) = E[X^2 + tX + k\mu_2]^2 \geq 0 \quad \forall \quad t,$$

where $\mu_r = EX^r$, ...(*i*)

is the $r$th moment of $X$ about mean.

$$\therefore \quad Z(t) = E[X^4 + t^2X^2 + k^2\mu_2^2 + 2tX^3 + 2k\mu_2X^2 + 2k\mu_2 tX]$$

$$= \mu_4 + t^2\mu_2 + k^2\mu_2^2 + 2t\mu_3 + 2k\mu_2^2$$

[Using (*i*) and $E(X) = 0$]

$$= t^2\mu_2 + 2t\mu_3 + \mu_4 + k^2\mu_2^2 + 2k\mu_2^2 \geq 0 \text{ for all } t. \quad \text{...(ii)}$$

Since $Z(t)$ is a quadratic form in $t$, $Z(t) \geq 0$ for all $t$ iff its discriminant is $\leq 0$, *i.e.*,

iff $$(2\mu_3)^2 - 4\mu_2[\mu_4 + k^2\mu_2^2 + 2k\mu_2^2] \leq 0$$

$$\Rightarrow \quad \frac{\mu_3^2}{\mu_2^3} - \left[\frac{\mu_4}{\mu_2^2} + k^2 + 2k\right] \leq 0 \quad \text{[Dividing by } 4\mu_2^3 > 0]$$

$$\Rightarrow \quad \beta_1 - \beta_2 - (2k + k^2) \leq 0$$
$$\Rightarrow \quad \beta_2 \geq \beta_1 - (2k + k^2)$$

**Deductions.** *(i)* Taking $k = 0$ in (*) we get $\beta_2 \geq \beta_1$ ...(**)

*(ii)* Taking $k = -1$ in (*) we get : $\beta_2 \geq \beta_1 + 1$ ...(***)

a result, which is established differently in Example 6·26.

*(iii)* Since $\beta_1 = \mu_3^2/\mu_2^3$ is always non-negative, we get from (***) :

$$\beta_2 \geq 1 \qquad ...(****)$$

**Remark.** The sign of equality holds in (****), *i.e.*, $\beta_2 = 1$ iff

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = 1 \quad \Rightarrow \quad \mu_4 = \mu_2^2$$

$$\Rightarrow \quad E[X - E(X)]^4 = \left[E(X - E(X))\right]^2$$
$$\Rightarrow \quad E(Y^2) - [E(Y)]^2 = 0, \qquad (Y = [X - E(X)]^2)$$
$$\Rightarrow \quad \text{Var}(Y) = 0$$
$$\Rightarrow \quad P[Y = E(Y)] = 1 \qquad [\textit{See Example } 6{\cdot}9]$$
$$\Rightarrow \quad P[(X - \mu)^2 = E(X - \mu)^2] = 1$$
$$\Rightarrow \quad P[(X - \mu)^2 = \sigma^2] = 1$$
$$\Rightarrow \quad P[(X - \mu) = \pm \sigma] = 1$$
$$\Rightarrow \quad P[X = \mu \pm \sigma] = 1$$

Thus $X$ takes only two values $\mu + \sigma$ and $\mu - \sigma$ with respective probabilities $p$ and $q$, (say).

$$\therefore \quad E(X) = p(\mu + \sigma) + q(\mu - \sigma) = \mu$$
$$\Rightarrow \quad p + q = 1 \quad \text{and} \quad (p - q)\sigma = 0$$

But since $\sigma \neq 0$, (since in this case $\beta_2$ is defined) we have :

$$p + q = 1 \quad \text{and} \quad p - q = 0 \quad \Rightarrow \quad p = q = 1/2.$$

Hence $\beta_2 = 1$ iff the r.v. $X$ assumes only two values, each with equal probability $1/2$.

**Example 6·25.** *Let X and Y be two variates having finite means.*

*Prove or disprove :*

*(a)* $E[\text{Min}(X, Y)] \leq \text{Min}[E(X), E(Y)]$

*(b)* $E[\text{Max}(X, Y)] \geq \text{Max}[E(X), E(Y)]$

*(c)* $E[\text{Min}(X, Y) + \text{Max}(X, Y)] = E(X) + E(Y)$

**[Delhi Univ. B.A. (Stat. Hons.), Spl. Course, 1989]**

**Solution.** We know that

$$\text{Min}(X, Y) = \frac{1}{2}(X + Y) - |X - Y| \qquad ...(i)$$

and

$$\text{Max}(X, Y) = \frac{1}{2}(X + Y) + |X - Y| \qquad ...(ii)$$

(a) $\therefore \quad E[\text{Min}(X, Y)] = \frac{1}{2} E(X+Y) - E|X-Y|$ ...(iii)

We have : $|E(X-Y)| \leq E|X-Y|$

$\Rightarrow \quad -|E(X-Y)| \geq -E|X-Y|$

$\Rightarrow \quad E|X-Y| \leq -|E(X-Y)| = -|E(X)-E(Y)|$ ...(*)

Substituting in (iii) we get :

$$E[\text{Min}(X, Y)] \leq \frac{1}{2} E(X+Y) - E|X-Y|$$

$$\leq \frac{1}{2}[E(X)+E(Y)] - |E(X)-E(Y)| \quad \text{...[From (*)]}$$

$\Rightarrow \quad E[\text{Min}(X, Y)] \leq \text{Min}[E(X), E(Y)]$

(b) Similarly from (ii) we get :

$$E[\text{Max}(X, Y)] = \frac{1}{2} E(X+Y) + E|X+Y|$$

$$\geq \frac{1}{2} E(X+Y) + |E(X+Y)|$$

$$(\because |E(X+Y)| \leq E|X+Y|)$$

$$= \frac{1}{2}[E(X)+E(Y)] + |E(X)+E(Y)|$$

$$= \text{Max}[E(X), E(Y)]$$

i.e., $E[\text{Max}(X, Y)] \geq \text{Max}[E(X), E(Y)]$

(c) $\quad [\text{Min}(X, Y) + \text{Max}(X, Y)] = [X+Y]$

$\Rightarrow \quad E[\text{Min}(X, Y) + \text{Max}(X, Y)] = E(X+Y)$

$$= E(X) + E(Y),$$

as required.

Hence all the results in (a), (b) and (c) are true.

**Example 6·26.** *Use the relation $E(AX^a + BX^b + CX^c)^2 \geq 0$, X being a random variable with $E(X) = 0$, E denoting the mathematical expectation, to show that*

$$\begin{vmatrix} \mu_{2a} & \mu_{a+b} & \mu_{a+c} \\ \mu_{a+b} & \mu_{2b} & \mu_{b+c} \\ \mu_{a+c} & \mu_{b+c} & \mu_{2c} \end{vmatrix} \geq 0, \quad \text{...(*)}$$

*$\mu_n$ denoting the nth moment about mean.*

*Hence or otherwise show that Pearson Beta- coefficients satisfy the inequality*

$$\beta_2 - \beta_1 - 1 \geq 0.$$

*Also deduce that $\beta_2 \geq 1$.*

**Solution.** Since $E(X) = 0$, we get

$$E(X^r) = \mu_r \quad \text{...(**)}$$

We are given that

$$E(AX^a + BX^b + CX^c)^2 \geq 0$$

$\Rightarrow \quad E[A^2X^{2a} + B^2X^{2b} + C^2X^{2c} + 2ABX^{a+b} + 2ACX^{a+c} + 2BCX^{b+c}] \geq 0$

$\Rightarrow \quad A^2\mu_{2a} + B^2\mu_{2b} + C^2\mu_{2c} + 2AB\mu_{a+b} + 2AC\mu_{a+c} + 2BC\mu_{b+c} \geq 0$

[From (**)] ...(***)

for all values of $A, B, C$.

We know from matrix theory that the conditions for the quadratic form

$$a'x^2 + b'y^2 + c'z^2 + 2f'yz + 2g'zx + 2h'xy,$$

to be non-negative for all values of $x$, $y$ and $z$ are

$$a' \geq 0, \quad (ii) \begin{vmatrix} a' & h' \\ h' & b' \end{vmatrix} \geq 0, \text{ and } \quad (iii) \begin{vmatrix} a' & h' & g' \\ h' & b' & f' \\ g' & f' & c' \end{vmatrix} \geq 0$$

Comparing with (***), we have

$$a' = \mu_{2a}, b' = \mu_{2b}, c' = \mu_{2c}, f' = \mu_{b+c}, g' = \mu_{a+c}, h' = \mu_{a+b}$$

Substituting these values in condition *(iii)* above, we get the required result.

Taking $a = 0, b = 1$ and $c = 2$ in (*) and noting that $\mu_0 = 1$ and $\mu_1 = 0$, we get

$$\begin{vmatrix} 1 & 0 & \mu_2 \\ 0 & \mu_2 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \end{vmatrix} \geq 0$$

$$\Rightarrow \qquad \mu_2 \mu_4 - \mu_3^2 + \mu_2 (-\mu_2^2) \geq 0$$

Dividing throughout by $\mu_2^3$ (assuming that $\mu_2$ is finite, for otherwise $\beta_2$ will become infinite), we get

$$\frac{\mu_4}{\mu_2^2} - \frac{\mu_3^2}{\mu_2^3} - 1 \geq 0$$

$$\Rightarrow \qquad \beta_2 - \beta_1 - 1 \geq 0$$

$$\Rightarrow \qquad \beta_2 \geq \beta_1 + 1$$

Further since $\beta_1 \geq 0$, we get $\beta_2 \geq 1$.

**Example 6·27.** *Let $X$ be a non-negative random variable with distribution function $F$. Show that*

$$E(X) = \int_0^\infty [1 - F(x)]\, dx. \qquad \ldots(i)$$

*Conjecture a corresponding expression for $E(X^2)$.*

**[Delhi Univ. M.Sc.(Stat). 1988]**

**Solution.** Since $X \geq 0$, we have :

$$\text{R.H.S.} = \int_0^\infty [1 - P(X \leq x)]\, dx = \int_0^\infty \left[1 - \int_0^x f(u)\, du\right] dx,$$

where $f(\cdot)$ is the p.d.f. of r.v. $X$.

$$\therefore \qquad \text{R.H.S.} = \int_0^\infty \left[\int_x^\infty f(u)\, du\right] dx; \qquad \ldots(ii)$$

From the integral in bracket *(ii)*, we have, $u \geq x$ and since $x$ ranges from 0 to $\infty$, $u$ also range from 0 to $\infty$. Further $u \geq x \Rightarrow x \leq u$ and since $x$ is non-negative, we have $0 \le x \le u$. [See Region $R_1$ in Remark 2 below]. Hence changing the order of integration in *(ii)*, [by Fubini's theorem for non-negative functions], we get

$$\text{R.H.S.} = \int_0^\infty \left[ \int_0^u 1 \cdot dx \right] f(u)\, du = \int_0^\infty u \cdot f(u)\, du$$

$$= E(X) \qquad \textit{[Since } f(\cdot) \textit{ is p.d.f. of } X\textit{]}$$

**Conjecture for** $E(X^2)$. Consider the integral :

$$\int_0^\infty 2x\,[1 - F(x)]\, dx = \int_0^\infty 2x \left( \int_x^\infty f(u)\, du \right) dx$$

$$= \int_0^\infty \left( \int_0^u 2x\, dx \right) f(u)\, du,$$

(By Fubini's theorem for non-negative functions).

$$= \int_0^\infty u^2 \cdot f(u) \cdot du = E(X^2)$$

**Remarks. 1.** If $X$ is a non-negative r.v. then

$$\text{Var}\, X = EX^2 - [E(X)]^2 = \int_0^\infty 2x\,[1 - F(x)]\, dx - \mu_x^2 \qquad ...(iii)$$

**2.** If we do not restrict ourselves to non-negative random variables only, we have the following more generalised result.

*If F denotes the distribution function of the random variable X then :*

$$E(X) = \int_0^\infty [1 - F(x)]\, dx - \int_{-\infty}^0 F(x)\, dx, \qquad ...(iv)$$

*provided the integrals exist finitely.*

*Proof of (iv).* The first integral has already been evaluated in the above example, *i.e.,*

$$\int_0^\infty [1 - F(x)]\, dx = \int_0^\infty u \cdot f(u) \cdot du \qquad ...(v)$$

Consider :

$$\int_{-\infty}^0 F(x)\, dx = \int_{-\infty}^0 P(X \le x)\, dx = \int_{-\infty}^0 \left( \int_{-\infty}^x f(u)\, du \right) dx$$

$$= \int_{-\infty}^0 \left( \int_u^0 1 \cdot dx \right) f(u)\, du$$

[Changing the order of integration in the Region $R_2$ where $u \le x$].

$$= \int_{-\infty}^{0} u \cdot f(u)\, du \qquad ...(vi)$$

Subtracting *(vi)* from *(v)*, we get :

$$\int_0^\infty [1 - F(x)]\, dx - \int_{-\infty}^0 F(x)\, dx = \int_0^\infty u f(u)\, du + \int_{-\infty}^0 u f(u)\, du$$

$$= \int_{-\infty}^{\infty} u f(u)\, du$$

$$= E(X), \qquad (\text{Since } f(\cdot) \text{ is p.d.f. of } X)$$

as desired.

In this generalised case,

$$\text{Var}(X) = \int_0^\infty 2x\,[1 - F_X(x) + F_X(-x)]\, dx - [E(X)]^2$$

3. The corresponding analogue of the above result for discrete random variable is given in the next Example 6·28.

**Example 6·28.** *If the possible values of a variate X are 0, 1, 2, 3, .... then*

$$E(X) = \sum_{n=0}^{\infty} P(X > n)$$

**[Delhi Univ. B.Sc. (Maths Hons.), 1987]**

**Solution.** Let $P(X = n) = p_n, \; n = 0, 1, 2, 3, \ldots \qquad ...(i)$

If $E(X)$ exists, then by definition :

$$E(X) = \sum_{n=0}^{\infty} n \,.\, P(X = n) = \sum_{n=1}^{\infty} n \,.\, p_n \qquad ...(ii)$$

Consider :

$$\sum_{n=0}^{\infty} P(X > n) = P(X > 0) + P(X > 1) + P(X > 2) + \ldots$$

$$= P(X \geq 1) + P(X \geq 2) + P(X \geq 3) + \ldots$$

$$= (p_1 + p_2 + p_3 + p_4 + \ldots\ldots)$$
$$+ (p_2 + p_3 + p_4 + \ldots\ldots)$$
$$+ (p_3 + p_4 + p_5 + \ldots\ldots)$$
$$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$= p_1 + 2p_2 + 3p_3 + \ldots\ldots$$

$$= \sum_{n=1}^{\infty} n\, p_n$$

$$= E(X) \qquad [\text{From } (ii)]$$

As an illustration of this result, see Problem 24 in Exercise 6(a).

**Aliter.** R.H.S. $= \sum_{n=0}^{\infty} P(X > n) = \sum_{n=1}^{\infty} P(X \geq n)$

$$= \sum_{n=1}^{\infty}\left(\sum_{x=n}^{\infty} p(x)\right)$$

Since the series is convergent and $p(x) \geq 0 \ \forall \ x$, by Fubini's theorem, changing the order of summation we get :

$$\text{R.H.S.} = \sum_{x=1}^{\infty}\left(\sum_{n=1}^{x} p(x)\right) = \sum_{x=1}^{\infty}\left\{p(x)\sum_{n=1}^{x} 1\right\}$$

since $x \geq n \Rightarrow n \leq x$ and $x$ assumes only positive integral values :

$$\therefore \quad \text{R.H.S.} = \sum_{x=1}^{\infty} x\,p(x) = \sum_{x=0}^{\infty} x\,p(x) = E(X)$$

**Example 6·29.** *For any variates X and Y, show that*

$$\left\{E(X+Y)^2\right\}^{1/2} \leq \left\{E(X^2)\right\}^{1/2} + \left\{E(Y^2)\right\}^{1/2} \qquad ...(*)$$

**Solution.** Squaring both sides in (*), we have to prove

$$E(X+Y)^2 \leq \left[\left\{E(X^2)\right\}^{1/2} + \left\{E(Y^2)\right\}^{1/2}\right]^2$$

$$\Rightarrow \quad E(X^2) + E(Y^2) + 2E(XY) \leq E(X^2) + E(Y^2) + 2\sqrt{E(X^2)\,E(Y^2)}$$

$$\Rightarrow \quad E(XY) \leq \sqrt{E(X^2)\,E(Y^2)}$$

$$\Rightarrow \quad [E(XY)]^2 \leq E(X^2)\,.\,E(Y^2).$$

This is nothing but Cauchy-Schwartz inequality. [For proof see Theorem 6·11 page 6·13.]

**Example 6·30.** *Let X and Y be independent non-degenerate variates. Prove that*

$$Var(XY) = Var(X)\,.\,Var(Y)$$

*iff* $\qquad E(X) = 0,\ E(Y) = 0$

**[Delhi Univ. B.Sc. (Maths Hons.), 1989]**

**Solution.** We have

$$\text{Var}(XY) = E(XY)^2 - [E(XY)]^2 = E(X^2Y^2) - [E(XY)]^2$$

$$= E(X^2)\,E(Y^2) - [E(X)]^2\,[E(Y)]^2, \qquad ...(*)$$

since $X$ and $Y$ are independent.

$$\left.\begin{array}{ll} \text{If} & E(X) = 0 = E(Y) \\ \text{then} & \text{Var}(X) = E(X^2) \text{ and } \text{Var}(Y) = E(Y^2) \end{array}\right\} \qquad ...(**)$$

Substituting from (**) in (*), we get

$$\text{Var}(XY) = \text{Var}(X)\,.\,\text{Var}(Y),$$

as desired.

**Only If.** We have to prove that if

$$\text{Var}(XY) = \text{Var}(X)\,.\,\text{Var}(Y) \qquad ...(***)$$

then $\qquad E(X) = 0$ and $E(Y) = 0$.

Now (***) gives, [on using (*)]

$$E(X^2)E(Y^2) - [E(X)]^2.[E(Y)]^2 = \left\{E(X^2) - [E(X)]^2\right\} \times \left\{E(Y^2) - [E(Y)]^2\right\}$$

$$= E(X^2)E(Y^2) - E(X^2)[E(Y)]^2 - [E(X)]^2 E(Y^2) + [E(X)]^2 [E(Y)]^2$$

$$\Rightarrow E(X^2).[E(Y)]^2 - [E(X)]^2.[E(Y)]^2 + [E(X)]^2 E(Y^2) - [E(X)]^2 [E(Y)]^2 = 0$$

$$\Rightarrow [E(Y)]^2 \{E(X^2) - [E(X)]^2\} + [E(X)]^2 \{E(Y^2) - [E(Y)]^2\} = 0$$

$$\Rightarrow \quad [E(Y)]^2 \text{Var}(X) + [E(X)]^2 \text{Var}(Y) = 0 \qquad \ldots(****)$$

Since each of the quantities $[E(X)]^2$, $[E(Y)]^2$, Var $(X)$ and Var $(Y)$ is non-negative and since $X$ and $Y$ are given to be non-degenerate random variables such that Var $(X) > 0$ and Var $(Y) > 0$, (****) holds only if we have $E(X) = 0 = E(Y)$, as required.

## EXERCISE 6(a)

1. *(a)* Define a random variable and its mathematical expectation.

*(b)* Show that the mathematical expectation of the sum of two random variables is the sum of their individual expectations and if two variables are independent, the mathematical expectation of their product is the product of their expectations.

Is the condition of independence necessary for the latter ? If not, what is the necessary condition ?

*(c)* If $X$ is a random variable, prove that $|E(X)| \leq E(|X|)$.

*(d)* If $X$ and $Y$ are two random variables such that $X \leq Y$, prove that

$$E(X) \leq E(Y).$$

*(e)* Prove that $E[(X - c)^2] = [\text{Var}(X)] + [E(X) - c]^2$, where $c$ is a constant.

2. Prove that

*(a)* $E(aX + bY) = a E(X) + b E(Y)$.

where $a$ and $b$ are any constants.

*(b)* $E(a) = a$, $a$ being a constant.

*(c)* $E[ag(X)] = aE[g(X)]$

*(d)* $E[g_1(X) + g_2(X) + \ldots + g_n(X)] = E[g_1(X)] + E[g_2(X)] + \ldots + E[g_n(X)]$

*(e)* $|E[g(X)]| \leq E[|g(X)|]$.

*(f)* If $g(X) \geq 0$, everywhere then $E[g(X)] \geq 0$.

*(g)* If $g(X) \geq 0$ everywhere and $E[g(X)] = 0$, then $g(X) = 0$, *i.e.*, the random variable $g(X)$ has a one point distribution at $X = 0$.

3. Show that if $X$ is non-negative random variable such that both $E(X)$ and $E(1/X)$ exist, then

$$E(1/X) \geq 1/E(X).$$

4. If $X$, $Y$ are independent random variables with $E(X) = \alpha$, $E(X^2) = \beta$, and $E(Y^k) = a_k$ ; $k = 1, 2, 3, 4$, find $E(XY + Y^2)^2$.

5. *(a)* If $X$ and $Y$ are two independent random variables, show that

$$\text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y).$$

(*b*) With usual notations, show that

$\text{Cov}(aX + bY, cX + dY) = ac\,\text{Var}(X) + bd\,\text{Var}(Y) + (ad + bc)\,\text{Cov}(X, Y)$

(*c*) Show that

$$\text{Cov}\left(\sum_{i=1}^{n} a_i X_i,\ \sum_{i=1}^{n} b_j X_j\right) = \sum_{j=1}^{n}\sum_{i=1}^{n} a_i b_j\,\text{Cov}(X_i, X_j)$$

6. (*a*) Define the indicator function $I_A(x)$ and show that $E(I_A(X)) = P(A)$.

(*b*) Prove that the probability function $P(X \in A)$ for set $A$ and the distribution function $F_X(x)$, $(-\infty < x < \infty)$, can be regarded as expectations of some random variable.

**Hint.** Define the indicator functions :

$$I_A(x) = 1 \text{ if } x \in A \qquad I_y(x) = 1 \text{ if } x \le y$$
$$\quad\;\; = 0 \text{ if } x \notin A \qquad\quad\;\; = 0 \text{ if } x > y$$

Then we shall get :

$E(I_A(X)] = P(X \in A)$ and $E(I_y(X)) = P(X \le y) = F_X(y)$

7. (*a*) Let $X$ be a continuous random variable with median $m$. Minimise $E\,|X - b|$, as an function of $b$.

**Ans.** $E\,|X - b|$ is minimum when $b = m =$ Median. This states that absolute sum of deviations of a given set of observations is minimum when taken about median. [See Example 5·19.]

(*b*) Let $X$ be a random variable such that $E\,|X| < \infty$. Show that $E\,|X - C|$ is minimised if we choose $C$ equal to the median of the distribution.

**[Delhi Univ. B.Sc. (Maths Hons.), 1988]**

8. If $X$ and $Y$ are symmetric, show that

$$E\left(\frac{X}{X+Y}\right) = \frac{1}{2}$$

**Hint.** $$1 = E\left[\frac{X+Y}{X+Y}\right] = E\left[\frac{X}{X+Y}\right] + E\left[\frac{Y}{X+Y}\right]$$

$$\Rightarrow \qquad 1 = 2E\left[\frac{X}{X+Y}\right] \qquad (\because X \text{ and } Y \text{ are symmetric.})$$

9. (*a*) If a r.v. $X$ has a symmetric density about the point '*a*' and if $E(X)$ exists, then

$$\text{Mean}(X) = \text{Median}(X) = a$$

**Hint.** Given $f(a - x) = f(a + x)$ ; $f(x)$ p.d.f. of $X$. prove that

$$E(X - a) = \int_{-\infty}^{\infty}(x - a)f(x)\,dx = \int_{-\infty}^{\infty}(x - a)f(x)\,dx + \int_{a}^{\infty}(x - a)f(x)\,dx = 0$$

(*b*) If $X$ and $Y$ are two random variables with finite variances, then show that

$$E^2(XY) \le E(X^2)\,.\,E(Y^2) \qquad \ldots(*)$$

When does the equality sign hold in (*) ? **[Indian Civil Service, 1987]**

**10.** Let $X$ be a non-negative arbitrary r.v. with distribution function $F$. show that

$$E(X) = \int_0^{\infty} [1 - F_X(x)]\, dx - \int_{-\infty}^{0} F_X(x)\, dx,$$

in the sense that, if either side exists, so does the other and the two are equal.

**[Delhi Univ. B.Sc. (Maths Hons.), 1992]**

**11.** Show that if $Y$ and $Z$ are independent random values of a variable $X$, the expected value of $(Y - Z)^2$ is twice the variance of the distribution of $X$.

**[Allahabad Univ. B.Sc., 1989]**

**Hint.** $E(Y) = E(Z) = E(X) = \mu$, (say) ; $\sigma_y^2 = \sigma_z^2 = \sigma_x^2 = \sigma^2$, (say). ...(*)

$$E(Y - Z)^2 = E(Y^2) + E(Z^2) - 2E(Y)E(Z)$$

( $\because Y, Z$ are independent )

$$= (\sigma_y^2 + \mu_y^2) + (\sigma_z^2 + \mu_z^2) - 2\mu^2$$

$$= 2\sigma^2 = 2\sigma_x^2$$ [On using (*)]

**12.** Given the following table :

| $x$ | $-3$ | $-2$ | $-1$ | $0$ | $1$ | $2$ | $3$ |
|---|---|---|---|---|---|---|---|
| $p(x)$ | 0·05 | 0·10 | 0·30 | 0 | 0·30 | 0·15 | 0·10 |

Compute (i) $E(X)$, (ii) $E(2X \pm 3)$, (iii) $E(4X + 5)$, (iv) $E(X^2)$

(v) $V(X)$, and (vi) $V(2X \pm 3)$.

**13.** (a) $A$ and $B$ throw with one die for a stake of Rs. 44 which is to be won by the playr who first throws a 6. If $A$ has the first throw, what are their respective expectations?

**Ans.** Rs. 24, Rs. 20.

(b) A contractor has to choose between two jobs. The first promises a profit of Rs. 1,20,000 with a probability of ¾ or a loss of Rs. 30,000 due to delays with a probability of ¼ ; the second promises a profit of Rs. 1,80,000 with a probability of ½ or a loss of Rs. 45,000 with a probability of ½. Which job should the contractor choose so as to maximise his expected profit?

(c) A random variable $X$ can assume any positive integral value $n$ with a probability porportional to $1/3^n$. Find the expectation of $X$.

**[Delhi Univ. B.Sc., Oct. 1987]**

**14.** Three tickets are chosen at random without replacement from 100 tickets numbered 1, 2, ..., 100. Find the mathematical expectation of the sum of the numbers on the tickets drawn.

**15.** (a) Three urns contain respectavily 3 green and 2 white balls, 5 green and 6 white balls and 2 green and 4 white balls. One ball is drawn from each urn. Find the expected number of white balls drawn out.

**Hint.** Let us define the r.v.

$X_i = 1$, if the ball drawn from $i$th urn is white
$= 0$; otherwise

Then the number of white balls drawn is $S = X_1 + X_2 + X_3$.

$$E(S) = E(X_1) + E(X_2) + E(X_3) = 1 \times \frac{2}{5} + 1 \times \frac{6}{11} + 1 \times \frac{4}{6} = \frac{266}{165}.$$

*(b)* Urn $A$ contains 5 cards numbered from 1 to 5 and urn $B$ contains 4 cards numbered from 1 to 4. One card is drawn from each of these urns. Find the probability function of the number which appears on the cards drawn and its mathematical expectation.

Ans. 11/4.

**16.** *(a)* Thirteen cards are drawn from its pack simultaneously. If the values of aces are 1, face cards 10 and others according to denomination, find the expectation of the total score in all the 13 cards.

**[Madurai Univ. B.Sc., Oct. 1990]**

*(b)* Let $X$ be a random variable with p.d.f. as given below :

| $x$ : | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p(x)$ : | 1/3 | 1/2 | 1/24 | 1/8 |

Find the expected value of $Y = (X - 1)^2$. **[Aligarh Univ. B.Sc. (Hons.), 1992]**

**17.** A player tosses 3 fair coins. He wins Rs. 8, if three heads occur; Rs. 3, if 2 heads occur and Re. 1, if one head occurs. If the game is to be fair, how much should he lose, if no heads occur? **[Punjab Univ. M.A. (Econ.), 1987]**

**Hint.** $X$ : Player's prize in Rs.

| $x$ | 8 | 3 | 1 | $a$ |
|---|---|---|---|---|
| No. of heads | 3 | 2 | 1 | 0 |
| $p(x)$ | 1/8 | 3/8 | 3/8 | 1/8 |

$E(X) = \Sigma x\, p(x) = \frac{1}{8}(8 + 9 + 3 + a)$

For the game to be fair, we have :

$E(X) = 0 \Rightarrow 20 + a = 0 \Rightarrow a = -20$

Hence the player loses Rs. 20, if no heads come up.

**18.** *(a)* A coin is tossed until a tail appears. What is the expectation of the number of tosses ?

Ans. 2.

*(b)* Find the expectation of *(i)* the sum, and *(ii)* the product, of number of points on $n$ dice when thrown.

Ans. *(i)* $7n/2$, *(ii)* $(7/2)^n$

**19.** *(a)* Two cards are drawn at random from ten cards numbered 1 to 10. Find the expectation of the sum of points on two cards.

*(b)* An urn contains $n$ cards marked from 1 to $n$. Two cards are drawn at a time. Find the mathematical expectation of the product of the numbers on the cards.

**[Mysore Univ. B.Sc., 1991]**

(c) In a lottery $m$ tickets are drawn out of $n$ tickets numbered from 1 to $n$. What is the expectation of the sum of the squares of numbers drawn ?

(d) A bag contains $n$ white and 2 black balls. Balls are drawn one by one without replacement until a black is drawn. If 0, 1, 2, 3, ... white balls are drawn before the first black, a man is to receive $0^2, 1^2, 2^2, 3^2, \ldots$ rupees respectively. Find his expectation. **[Rajasthan Univ. B.Sc., 1992]**

(e) Find the expectation and variance of the number of successes in a series of independent trials, the probability of success in the $i$th trial being $p_i$ $(i = 1, 2, \ldots, n)$. **[Nagarjuna Univ. B.Sc., 1991]**

**20.** Balls are taken one by one out of an urn containing $w$ white and $b$ black balls until the first white ball is drawn. Prove that the expectation of the number of black balls preceding the first white ball is $b/(w+1)$.

**[Allahabad Univ. B.Sc. (Hons.), 1992]**

**21.** (a) $X$ and $Y$ are independent variables with means 10 and 20, and variances 2 and 3 respectively. Find the variance of $3X + 4Y$.

**Ans.** 66.

(b) Obtain the variance of $Y = 2X_1 + 3X_2 + 4X_3$ where $X_1, X_2$ and $X_3$ are three random variables with means given by 3, 4, 5 respectively, variances by 10, 20, 30 respectively, and co-variances by $\sigma_{X_1X_2} = 0$, $\sigma_{X_2X_3} = 0$, $\sigma_{X_1X_3} = 5$, where $\sigma_{X_rX_s}$ stands for the co-variance of $X_r$ and $X_s$.

**22.** (a) Suppose that $X$ is a random variable for which $E(X) = 10$ and $\text{Var}(X) = 25$. Find the positive values of $a$ and $b$ such that $Y = aX - b$, has expectation 0 and variance 1.

**Ans.** $a = 1/5$, $b = 2$

(b) Let $X_1$ and $X_2$ be two stochastic random variables having variances $k$ and 2 respectively. If the variance of $Y = 3X_2 - X_1$ is 25, find $k$.

**(Poona Univ. B.Sc., 1990)**

**Ans.** $k = 7$.

**23.** A bag contains $2n$ counters, of which half are marked with odd numbers and half with even numbers, the sum of all the numbers being $S$. A man is to draw two counters. If the sum of the numbers drawn is odd, he is to receive that number of rupees, if even he is to pay that number of rupees. Show that his expectation is $S/[n(2n-1)]$ rupees. **(I.F.S., 1989)**

**24.** A jar has $n$ chips numbred 1, 2, ..., $n$. A person draws a chip, returns it, draws another, returns it, and so on, until a chip is drawn that has been drawn before. Let $X$ be the number of drawings. Find $E(X)$.

**[Delhi Univ. B.A. (Stat. Hons.), Spl. Course, 1986]**

**Hint.** Obviously $P(X > 1) = 1$, because we must have at least two draws to get the chip which has been drawn before.

$$P(X > r) = P\,[\text{Distinct number on } i\text{th draw};\ i = 1, 2, \ldots, r\,]$$

$$= \frac{n}{n} \times \frac{n-1}{n} \times \frac{n-2}{n} \times \dots \times \frac{n-(r-1)}{n}$$

$$P(X > r) \Rightarrow \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\dots\dots\left(1 - \frac{r-1}{n}\right); r = 1, 2, 3, \dots \qquad \dots(*)$$

Hence, using the result in Example 6·28

$$E(X) = \sum_{r=0}^{\infty} P(X > r)$$

$$= P(X > 0) + P(X > 1) + P(X > 2) + \dots$$

$$= 1 + 1 + \left(1 - \frac{1}{n}\right) + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) +$$

$$\dots + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\dots\left(1 - \frac{r-1}{n}\right) + \dots \qquad \text{[Using (*)]}$$

**25.** A coin is tossed four times. Let $X$ denote the number of times a head is followed immediately by a tail. Find the distribution, mean and variance of $X$.

**Hint.** $S = \{H, T\} \times \{H, t\} \times \{H, T\} \times \{H, T\}$

$= \{HHHH, HHHT, HHTH, HTHH, HTHT, \dots\dots, TTTT\}$

$X$: 0, 1, 1, 1, 2, ..., 0

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $p(x)$ | 5/16 | 10/16 | 1/16 |

$E(X) = 3/4$, $E(X^2) = 7/8$

Var $X = 7/8 - 9/16 = 5/16$.

**26.** An urn contains balls numbered 1, 2, 3. First a ball is drawn from the urn and then a fair coin is tossed the number of times as the number shown on the drawn ball. Find the expected number of heads.

**[Delhi Univ. B.Sc. (Maths Hons.), 1984]**

**Hint.** $B_j$ : Event of drawing the ball numbered $j$.

$P(B_j) = 1/3$ ; $j = 1, 2, 3$.

$X$ : No. of heads shown. $X$ is a r.v. taking the values 0, 1, 2, and 3.

$$P(X = x) = \sum_{j=1}^{3} P(B_j) \cdot P(X = x \mid B_j) = \frac{1}{3} \sum_{j=1}^{3} P(X = x \mid B_j)$$

$$\therefore P(X = 0) = \frac{1}{3}[P(X = 0 \mid B_1) + P(X = 0 \mid B_2) + P(X = 0 \mid B_3)]$$

$$= \frac{1}{3}\left[\frac{1}{2} + \frac{1}{4} + \frac{1}{8}\right] = \frac{7}{24}$$

$$P(X = 1) = \frac{1}{3}[P(X = 1 \mid B_1) + P(X = 1 \mid B_2) + P(X = 1 \mid B_3)]$$

$$= \frac{1}{3}\left[\frac{1}{2} + \frac{2}{4} + \frac{3}{8}\right]$$

*e.g.*, $P(X = 0 \mid B_2) = P$ [ No head when two coins are tossed ] $= 2/4$

$P(X = 1 \mid B_3) = P$ [ 1 head when three coins are tossed ] $= 3/8$

Similarly $P(X = 2) = \frac{1}{3}\left(0 + \frac{1}{4} + \frac{3}{8}\right) = \frac{5}{24}$

$$P(X=3)=\frac{1}{3}\left(0+0+\frac{1}{8}\right)=\frac{1}{24}$$

$$\therefore \quad E(X)=\sum_{x=0}^{3} x\, P(X=x)=\frac{11}{24}+\frac{10}{24}+\frac{3}{24}=1$$

**27.** An urn contains $pN$ white and $qN$ black balls, the total number of balls being $N$, $p+q=1$. Balls are drawn one by one (without being returned to the urn) until a certain number $n$ of balls is reached.

Let $X_i = 1$, if the $i$th ball drawn is white.
$= 0$, if the $i$th ball drawn is black.

(*i*) Show that $E(X_i)=p$, $\text{Var}(X_i)=pq$.

(*ii*) Show that the co-variance between $X_j$ and $X_k$ is $-\dfrac{pq}{n-1}$, $(j \neq k)$

(*iii*) From (*i*) and (*ii*), obtain the variance of $S_n = X_1 + X_2 + \ldots + X_n$.

**28.** Two similar decks of $n$ distinct cards each are put into random order and are matched against each other. Prove that the probability of having exactly $r$ matches is given by

$$\frac{1}{r\,!}\sum_{k=0}^{n-r}\frac{(-1)^k}{k\,!},\quad r=0,1,2,\ldots n$$

Prove further that the expected number of matches and its variance are equal and are independent of $n$.

**29.** (*a*) If $X$ and $Y$ are two independent random variables, such that $E(X)=\lambda_1$, $V(X)=\sigma_1^2$ and $E(Y)=\lambda_2$, $V(Y)=\sigma_2^2$, then prove that

$$V(XY)=\sigma_1^2\sigma_2^2+\lambda_1^2\sigma_2^2+\lambda_2^2\sigma_1^2$$ **[Gorakhpur Univ. B.Sc., 1992]**

(*b*) If $X$ and $Y$ are two independent random variables, show that

$$\frac{V(XY)}{[E(X)]^2[E(Y)]^2}=C_X^2C_Y^2+C_X^2+C_Y^2$$

where $$C_X=\frac{\sqrt{V(X)}}{E(X)},\quad C_Y=\frac{\sqrt{V(Y)}}{E(Y)}$$

are the so-called coefficients of variation of $X$ and $Y$? **[Patna Univ. B.Sc., 1991]**

**30.** A point $P$ is taken at random in a line $AB$ of length $2a$, all positions of the point being equally likely. Show that the expected value of the area of the rectangle $AP \cdot PB$ is $2a^2/3$ and the probability of the area exceeding $1/2a^2$ is $1/\sqrt{2}$. **[Delhi Univ. B.Sc. (Maths Hons.), 1986]**

**31.** If $X$ is a random variable with $E(X)=\mu$ satisfying $P(X\leq 0)=0$, show that $P(X>2\mu)\leq 1/2$. **[Delhi Univ. B.Sc. (Maths Hons.), 1992]**

## OBJECTIVE TYPE QUESTIONS

**1.** Fill in the blanks :

(*i*) Expected value of a random variable $X$ exists if .........

(*ii*) If $E(X^r)$ exists then $E(X^s)$ also exists for .........

(*iii*) When $X$ is a random variable, expectation of $(X-\text{constant})^2$ is minimum when the constant is ....

(iv) $E\,|X - A|$ is minimum when $A =$ ....

(v) Var $(c) =$ ..., where $c$ is a constant

(vi) Var $(X + c) =$ ..., where $c$ is a constant

(vii) Var $(aX + b) =$ ..., where a and b are constants.

(viii) If $X$ is a r.v. with mean $\mu$ and variance $\sigma^2$ then

$$E\left(\frac{X-\mu}{\sigma}\right) = ...., \text{ Var}\left(\frac{X-\mu}{\sigma}\right) = ....$$

(ix) $[E(XY)]^2$ .... $E(X^2) \cdot E(Y^2)$.

(x) $V(aX \pm bY) = \ldots$

where $a$ and $b$ are constants.

**II.** Mark the correct answer in the following :

(i) For two random variables $X$ and $Y$, the relation

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

holds good

(a) if $X$ and $Y$ are statistically independent,

(b) for all $X$ and $Y$,

(c) if $X$ and $Y$ are identical.

(ii) Var $(2X \pm 3)$ is

(a) 5 (b) 13 (c) 4, if Var $X = 1$.

(iii) $E(X - k)^2$ is minimum when

(a) $k < E(X)$, (b) $k > E(X)$, (c) $k = E(X)$.

**III.** Comment on the following :

If $X$ and $Y$ are mutually independent variables, then

(i) $E(XY + Y + 1) - E(X + 1)\,E(Y) = 0$

(ii) $X$ and $Y$ are independent if and only if

$$\text{Cov}(X, Y) = 0$$

(iii) For every univariable distribution :

(a) $V(cX) = c^2 V(X)$ (b) $E(c/X) = c/E(X)$

(iv) Expected value of a r.v. always exists.

**IV.** Mark true or false with reasons for your answers :

(a) Cov $(X, Y) = 0 \Rightarrow X$ and $Y$ are independent.

(b) If Var $(X) >$ Var $(Y)$, then $X + Y$ and $X - Y$ are dependent.

(c) If Var $(X) =$ Var $(Y)$ and if $2X + Y$ and $X - Y$ are independent, then $X$ and $Y$ are dependent.

(d) If Cov $(aX + bY, bX + aY) \neq ab$ Var $(X + Y)$, then $X$ and $Y$ are dependent.

**6·8. Moments of Bivariate Probability Distributions.** The mathematical expectation of a function $g(x, y)$ of two-dimensional random variable $(X, Y)$ with p.d.f. $f(x, y)$ is given by

$$E[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx\, dy \qquad \text{...(6·43)}$$

(If $X$ and $Y$ are continuous variables)

$$= \sum_i \sum_j x_i y_j P(X = x_i \cap Y = y_j), \qquad \text{...(6·43 } a)$$

(If $X$ and $Y$ are discrete variables)

provided the expectation exists.

In particular, the $r$th and $s$th product moment about origin of the random variables $X$ and $Y$ respectively is defined as

$$\mu_{rs}' = E(X^r Y^s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^r y^s f(x, y)\, dx\, dy$$

or $\quad \mu_{rs}' = \sum_i \sum_j x_i^r y_j^s \, P(X = x_i \cap Y = y_j) \qquad \text{...(6·44)}$

The joint $r$th central moment of $X$ and $s$th central moment of $Y$ is given by

$$\mu_{rs} = E\left[\{X - E(X)\}^r \{Y - E(Y)\}^s\right]$$

$$= E\left[(X - \mu_X)^r (Y - \mu_Y)^s\right], \quad [E(X) = \mu_X, E(Y) = \mu_Y] \qquad \text{...(6·45)}$$

In particular

$$\mu_{00}' = 1 = \mu_{00}\ , \quad \mu_{10} = 0 = \mu_{01}$$
$$\mu_{10}' = E(X)\ , \quad \mu_{01}' = E(Y)$$
$$\mu_{20} = \sigma_X^2\ , \quad \mu_{02} = \sigma_Y^2 \text{ and } \mu_{11} = \text{Cov}(X, Y).$$

**6·9. Conditional Expectation and Conditional Variance.**

**Discrete Case.** The conditional expectation or mean value of a continuous function $g(X, Y)$ given that $Y = y_j$, is defined by

$$E\{g(X, Y) \mid Y = y_j\} = \sum_{i=1}^{\infty} g(x_i, y_j)\ P(X = x_i \mid Y = y_j)$$

$$= \frac{\sum_{i=1}^{\infty} g(x_i, y_j)\ P(X = x_i \cap Y = y_j)}{P(Y = y_j)} \qquad \text{...(6·46)}$$

*i.e.*, $E[g(X, Y) \mid Y = y_j]$ is nothing but the expectation of the function $g(X, y_j)$ of $X$ w.r.t. the conditional distribution of $X$ when $Y = y_j$.

In particular, the conditional expectation of a discrete random variable $X$ given $Y = y_j$ is

$$E(X \mid Y = y_j) = \sum_{i=1}^{\infty} x_i\ P(X = x_i \mid Y = y_j) \qquad \text{...(6·47)}$$

The conditional variance of $X$ given $Y = y_j$ is likewise given by

$$V(X \mid Y = y_j) = E[\{X - E(X \mid Y = y_j)\}^2 \mid Y = y_j] \qquad \text{...(6·47a)}$$

The conditional expectation of $g(X, Y)$ and the conditional variance of $Y$ given $X = x_i$ may also be defined in an exactly similar manner.

**Continuous Case.** The conditional expectation of $g(X, Y)$ on the hypothesis $Y = y$ is defined by

$$E\{g(X, Y) \mid Y = y\} = \int_{-\infty}^{\infty} g(x, y) f_{X \mid Y}(x \mid y)\, dx$$

$$= \frac{\int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx}{f_Y(y)} \qquad \ldots(6\cdot48)$$

In particular, the conditional mean of $X$ given $Y = y$ is defined by

$$E(X \mid Y = y) = \frac{\int_{-\infty}^{\infty} x f(x, y)\, dx}{f_Y(y)}$$

Similarly, we define

$$E(Y \mid X = x) = \frac{\int_{-\infty}^{\infty} y f(x, y)\, dy}{f_X(x)} \qquad \ldots(6\cdot48\ a)$$

The conditional variance of $X$ may be defined as

$$V(X \mid Y = y) = E\left[\{X - E(X \mid Y = y)\}^2 \mid Y = y\right]$$

Similarly, we define

$$V(Y \mid X = x) = E\left[\{Y - E(Y \mid X = x)\}^2 \mid X = x\right] \qquad \ldots(6\cdot49)$$

**Theorem 6·13.** *The expected value of X is equal to the expectation of the conditional expectation of X given Y. Symbolically,*

$$E(X) = E[E(X \mid Y)] \qquad \ldots(6\cdot50)$$

**[Calicut Univ. B.Sc. (Main Stat.), 1980]**

**Proof.** $E[E(X \mid Y)] = E[\sum_i x_i\ P(X = x_i \mid Y = y_j)]$

$$= E\left[\sum_i x_i \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}\right]$$

$$= \sum_j \left[\sum_i \left\{x_i \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}\right\}\right] P(Y = y_j)$$

$$= \sum_j \sum_i x_i \cdot P(X = x_i \cap Y = y_j)$$

$$= \sum_i \left[x_i \{\sum_j P(X = x_i \cap Y = y_j)\}\right]$$

$$= \sum x_i P(X = x_i) = E(X).$$

**Theorem 6·14.** *The variance of X can be regarded as consisting of two parts, the expectation of the conditional variance and the variance of the conditional expectation.* Symbolically,

$$V(X) = E[V(X \mid Y)] + V[E(X \mid Y)] \qquad \ldots(6\cdot51)$$

**Proof.** $E[V(X \mid Y)] + V[E(X \mid Y)]$

$$= E\Big[E(X^2 \mid Y) - \{E(X \mid Y)\}^2\Big] + E[\{E(X \mid Y)\}^2] - [E\{E(X \mid Y)\}]^2$$

$$= E[E(X^2 \mid Y)] - E[\{E(X \mid Y)\}^2] + E[\{E(X \mid Y)\}^2] - [E\{E(X \mid Y)\}]^2$$

$$= E[E(X^2 \mid Y)] - [E(X)]^2 \qquad (c.f.\ \text{Theorem } 6\cdot13)$$

$$= E[\sum_i x_i^2 P(X = x_i \mid Y = y_j)] - [E(X)]^2$$

$$= E\left[\sum_i x_i^2 \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}\right] - [E(X)]^2$$

$$= \sum_j \left\{\left[\sum_i x_i^2 \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}\right] P(Y = y_j)\right\} - [E(X)]^2$$

$$= \sum_i \left[x_i^2 \sum_j P(X = x_i \cap Y = y_j)\right] - [E(X)]^2$$

$$= \sum_i x_i^2 P(X = x_i) - [E(X)]^2$$

$$= E(X^2) - [E(X)]^2 = \text{Var}(X)$$

Hence the theorem.

**Remarks** The proofs of Theorems 6·13 and 6·14 for continous r.v.'s X and Y are left as an exercise to the reader.

**Theorem 6·15.** *Let A and B be two mutually exclusive events, then*

$$E(X \mid A \cup B) = \frac{P(A)\,E(X \mid A) + P(B)\,E(X \mid B)}{P(A \cup B)} \qquad \ldots(6\cdot52)$$

*where by def.,*

$$E(X \mid A) = \frac{1}{P(A)} \sum_{x_i \in A} x_i P(X = x_i)$$

**Proof.** $E(X \mid A \cup B) = \dfrac{1}{P(A \cup B)} \sum\limits_{x_i \in A \cup B} x_i\, P(X = x_i)$

Since $A$ and $B$ are mutually exclusive events,

$$\sum_{x_i \in A \cup B} x_i P(X = x_i) = \sum_{x_i \in A} x_i P(X = x_i) + \sum_{x_i \in B} x_i P(X = x_i)$$

$$\therefore\ E(X \mid A \cup B) = \frac{1}{P(A \cup B)} [P(A)\,E(X \mid A) + P(B)\,E(X \mid B)]$$

**Cor.** $E(X) = P(A)\,E(X \mid A) + P(\overline{A})\,E(X \mid \overline{A}) \qquad \ldots(6\cdot53)$

The corollary follows by putting $B = \overline{A}$ in the above Theorem.

**Example 6·31.** *Two ideal dice are thrown. Let $X_1$ be the score on the first die and $X_2$ the score on the second die. Let Y denote the maximum of $X_1$ and $X_2$, i.e., $Y = \max(X_1, X_2)$.*

*(i) Write down the joint distribution of Y and $X_1$,*

*(ii) Find the mean and variance of Y and co-variance $(Y, X_1)$.*

**Solution.** Each of the random variables $X_1$ and $X_2$ can take six values 1, 2, 3, 4, 5, 6 each with probability 1/6, *i.e.,*

$$P(X_1 = i) = P(X_2 = i) = 1/6 \;;\; i = 1, 2, 3, 4, 5, 6 \qquad ...(i)$$

$$Y = \text{Max}(X_1, X_2).$$

Obviously

$$P(X_1 = i, Y = j) = 0\,, \text{ if } j < i = 1, 2, ..., 6$$

$$P(X_1 = i, Y = i) = P(X_1 = i, X_2 \le i) = \sum_{j=1}^{i} P(X_1 = i, X_2 = j)$$

$$= \sum_{j=1}^{i} P(X_1 = i)\,P(X_2 = j) \qquad (\because X_1, X_2 \text{ are independent.})$$

$$= \sum_{j=1}^{i} \left(\frac{1}{36}\right) = \frac{i}{36} \;;\; i = 1, 2, ..., 6.$$

$$P(X_1 = i, Y = j) = P(X_1 = i, X_2 = j) \;;\; j > i$$

$$= P(X_1 = i)\,P(X_2 = j) = \frac{1}{36} \;;\; j > i = 1, 2, ..., 6.$$

The joint probability table of $X_1$ and $Y$ is given as follows:

| $X_1$ \ $Y$ | 1 | 2 | 3 | 4 | 5 | 6 | *Marginal Totals* |
|---|---|---|---|---|---|---|---|
| 1 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 6/36 |
| 2 | 0 | 2/36 | 1/36 | 1/36 | 1/36 | 1/36 | 6/36 |
| 3 | 0 | 0 | 3/36 | 1/36 | 1/36 | 1/36 | 6/36 |
| 4 | 0 | 0 | 0 | 4/36 | 1/36 | 1/36 | 6/36 |
| 5 | 0 | 0 | 0 | 0 | 5/36 | 1/36 | 6/36 |
| 6 | 0 | 0 | 0 | 0 | 0 | 6/36 | 6/36 |
| *Marginal Totals* | 1/36 | 3/36 | 5/36 | 7/36 | 9/36 | 11/36 | 1 |

$$E(Y) = 1.\frac{1}{36} + 2.\frac{3}{36} + 3.\frac{5}{36} + 4.\frac{7}{36} + 5.\frac{9}{36} + 6.\frac{11}{36}$$

$$= \frac{1}{36}\left[1 + 6 + 15 + 28 + 45 + 66\right] = \frac{161}{36}$$

$$E(Y^2) = 1^2.\frac{1}{36} + 2^2.\frac{3}{36} + 3^2.\frac{5}{36} + 4^2.\frac{7}{36} + 5^2.\frac{9}{36} + 6^2.\frac{11}{36} = \frac{791}{36}$$

$$V(Y) = E(Y^2) - [E(Y)]^2 = \frac{791}{36} - \left(\frac{161}{36}\right)^2 = \frac{2555}{1296}$$

$$E(X_1) = \frac{6}{36}\left[1+2+3+4+5+6\right] = \frac{126}{36} = \frac{21}{6}$$

$$E(X_1 Y) = 1.\frac{1}{36} + 2.\frac{1}{36} + 3.\frac{1}{36} + 4.\frac{1}{36} + 5.\frac{1}{36} + 6.\frac{1}{36}$$
$$+ 4.\frac{2}{36} + 6.\frac{1}{36} + 8.\frac{1}{36} + 10.\frac{1}{36} + 12.\frac{1}{36}$$
$$+ 9.\frac{3}{36} + 12.\frac{1}{36} + 15.\frac{1}{36} + 18.\frac{1}{36}$$
$$+ 16.\frac{4}{36} + 20.\frac{1}{36} + 24.\frac{1}{36}$$
$$+ 25.\frac{5}{36} + 30.\frac{1}{36} + 36.\frac{6}{36}$$
$$= \frac{1}{36}\left[21 + 44 + 72 + 108 + 155 + 216\right] = \frac{1}{36} \times 616$$

$$\text{Cov}(X_1, Y) = E(X_1 Y) - E(X_1) E(Y)$$
$$= \frac{616}{36} - \frac{21}{6} \cdot \frac{161}{36} = \frac{3696 - 3381}{216} = \frac{315}{216}.$$

**Example 6·32.** *Let X and Y be two random variables each taking three values −1, 0 and 1, and having the joint probability distribution :*

*(i) Show that X and Y have different expectations.*

| Y \ X | −1 | 0 | 1 | *Total* |
|---|---|---|---|---|
| −1 | 0 | ·1 | ·1 | ·2 |
| 0 | ·2 | ·2 | ·2 | ·6 |
| 1 | 0 | ·1 | ·1 | ·2 |
| *Total* | 2 | ·4 | ·4 | 1 0 |

*(ii) Prove that X and Y are uncorrelated.*

*(iii) Find Var X and Var Y.*

*(iv) Given that Y = 0, what is the conditional probability distribution of X ?*

*(v) Find $V(Y \mid X = -1)$.*

**Solution.** *(i)* $E(Y) = \Sigma p_i y_i = -1(\cdot 2) + 0(\cdot 6) + 1(\cdot 2) = 0$

$$E(X) = \Sigma p_i x_i = -1(\cdot 2) + 0(\cdot 4) + 1(\cdot 4) = \cdot 2$$

$$\therefore \quad E(X) \neq E(Y)$$

*(ii)* $E(XY) = \Sigma p_{ij} y_i x_j$

$$= (-1)(-1)(0) + 0(-1)(\cdot 1) + 1(-1)(\cdot 1)$$
$$+ 0(-1)(\cdot 2) + 0(0)(\cdot 2) + 0(1)(\cdot 2)$$
$$+ 1(-1)(0) + 1(0)(\cdot 1) + 1(1)(\cdot 1)$$
$$= -0\cdot 1 + 0\cdot 1 = 0$$

$$\therefore \quad \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$$

$\Rightarrow$ $X$ and $Y$ are uncorrelated (*c.f.* § 10·3)

(iii) $E(Y^2) = (-1)^2(\cdot 2) + 0(\cdot 6) + 1^2(\cdot 2) = \cdot 4$

$\therefore$ $V(Y) = E(Y^2) - [E(Y)]^2 = \cdot 4$

$E(X^2) = (-1)^2(\cdot 2) + 0(\cdot 4) + 1^2(\cdot 4) = \cdot 2 + \cdot 4 = \cdot 6$

$V(X) = \cdot 6 - \cdot 04 = \cdot 56$

(iv) $$P(X = -1 \mid Y = 0) = \frac{P(X = -1 \cap Y = 0)}{P(Y = 0)} = \frac{\cdot 2}{\cdot 6} = \frac{1}{3}$$

$$P(X = 0 \mid Y = 0) = \frac{P(X = 0 \cap Y = 0)}{P(Y = 0)} = \frac{\cdot 2}{\cdot 6} = \frac{1}{3}$$

$$P(X = 1 \mid Y = 0) = \frac{P(X = 1 \cap Y = 0)}{P(Y = 0)} = \frac{\cdot 2}{\cdot 6} = \frac{1}{3}$$

(v) $$V(Y \mid X = -1) = E(Y \mid X = -1)^2 - \{E(Y \mid X = -1)\}^2$$

$$E(Y \mid X = -1) = \sum_y y\, P(Y = y \mid X = -1) = (-1)0 + 0(\cdot 2) + 1(0) = 0$$

$$E(Y \mid X = -1)^2 = \sum_y y^2\, P(Y = y \mid X = -1) = 1(0) + 0(\cdot 2) + 1(0) = 0$$

$\therefore$ $V(Y \mid X = -1) = 0.$

**Example 6·33.** *Two tetrahedra with sides numbered 1 to 4 are tossed. Let $X$ denote the number on the downturned face of the first tetrahedron and $Y$ denote the larger of the downturned numbers. Investigate the following :*

*(a) Joint density function of $X$, $Y$ and marginals $f_X$ and $f_Y$,*

*(b) $P\{X \leq 2, Y \leq 3\}$,* *(c) $\rho(X, Y)$,* *(d) $E(Y \mid X = 2)$,*

*(e) Construct joint density different from that in part (a) but possessing same marginals $f_X$ and $f_Y$.* **[Delhi Univ. B.A. (Stat. Hons.), Spl. Course, 1985]**

**Hint.** The sample space is $S = \{1, 2, 3, 4\} \times \{1, 2, 3, 4\}$ and each of the 16 sample points (outcomes) has probability $p = 1/16$ of occurrence.

Let $X$ : Number on the first dice and $Y$ : Larger of the numbers on the two dice. Then the above 16 sample points, in that order, give the following distribution of $X$ and $Y$.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Sample Point* | : | (1, 1) | (1, 2) | (1, 3) | (1, 4) | (2, 1) | (2, 2) | (2, 3) | (2, 4) |
| $X$ | : | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 |
| $Y$ | : | 1 | 2 | 3 | 4 | 2 | 2 | 3 | 4 |
| *Sample Point* | : | (3, 1) | (3, 2) | (3, 3) | (3, 4) | (4, 1) | (4, 2) | (4, 3) | (4, 4) |
| $X$ | : | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 |
| $Y$ | : | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 |

Since each sample point has probability $p = 1/16$, the joint density functions of $X$ and $Y$ and the marginal densities $f_X$ and $f_Y$ are given on page 6·61.

Here $p = 1/16$.

(b) $P(X \leq 2, Y \leq 3) = p + p + 2p + p + p = 6p = 3/8.$

(c) $\text{Var}(X) = EX^2 - [E(X)]^2 = \frac{15}{2} - \frac{25}{4} = \frac{5}{4}$ (Try it)

| | (a) $x$: 1 | 2 | 3 | 4 | Total $(f_y)$ | | (e) $x$: 1 | 2 | 3 | 4 | Total $(f_y)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $y$ 1 | $p$ | 0 | 0 | 0 | $p$ | $y$ 1 | $p$ | 0 | 0 | 0 | $p$ |
| 2 | $p$ | $2p$ | 0 | 0 | $3p$ | 2 | $p$ | $2p$ | 0 | 0 | $3p$ |
| 3 | $p$ | $p$ | $3p$ | 0 | $5p$ | 3 | $p$ | $p+\epsilon$ | $3p-\epsilon$ | 0 | $5p$ |
| 4 | $p$ | $p$ | $p$ | $4p$ | $7p$ | 4 | $p$ | $p-\epsilon$ | $p+\epsilon$ | $4p$ | $7p$ |
| Total $(f_x)$ | $4p$ | $4p$ | $4p$ | $4p$ | $16p=1$ | Total $(f_x)$ | $4p$ | $4p$ | $4p$ | $4p$ | 1 |

$$\text{Var}(Y) = EY^2 - [E(Y)]^2 = \frac{85}{8} - \left(\frac{25}{8}\right)^2 = \frac{55}{64} \qquad \text{(Try it)}$$

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{135}{16} - \frac{5}{2} \times \frac{25}{8} = \frac{5}{8} \qquad \text{(Try it)}$$

$$\therefore \quad \rho(X, Y) = \frac{5/8}{\sqrt{5/4 \times 55/64}} = \frac{2}{\sqrt{11}}$$

*(d)* $$E(Y \mid X = 2) = \Sigma y \,.\, f(y \mid x = 2) = \Sigma y \,.\, \frac{f(x = 2 \cap y)}{f(x = 2)}$$

$$= 4 \,.\, \Sigma y f(2, y) = 4\,[0 + 4p + 3p + 4p] = 44p = \frac{11}{4}$$

*(e)* Let $0 < \epsilon < p$. The joint density of $X$ and $Y$ given in (*e*) above is different from that in (*a*) but has the same marginals as in (*a*).

**Example 6·34.** *(a) Given two variates $X_1$ and $X_2$ with joint density function $f(x_1, x_2)$, prove that conditional mean of $X_2$ (given $X_1$) coincides with (unconditional) mean only if $X_1$ and $X_2$ are independent (stochastically).*

*(b) Let $f(x_1, x_2) = 21x_1^2 x_2^3$, $0 < x_1 < x_2 < 1$, and zero elsewhere be the joint p.d.f. of $X_1$ and $X_2$. Find the conditional mean and variance of $X_1$ given $X_2 = x_2$, $0 < x_2 < 1$.* **[Delhi Univ. M.A. (Eco.), 1986]**

**Solution.** *(a)* Conditional mean of $X_2$ given $X_1$ is given by :

$$E(X_2 \mid X_1 = x_1) = \int_{x_2} x_2 f(x_2 \mid x_1)\, dx_2 \qquad \ldots(*)$$

where $f(x_2 \mid x_1)$ is conditional p.d.f. of $X_2$ given $X_1 = x_1$.

But the joint p.d.f. of $X_1$ and $X_2$ is given by

$$f(x_1, x_2) = f_1(x_1) \,.\, f(x_2 \mid x_1)$$

$$\Rightarrow \qquad f(x_2 \mid x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}$$

where $f_1(\cdot)$ is marginal p.d.f. of $X_1$.

Substituting in (*), we get

$$E(X_2 \mid X_1 = x_1) = \int_{x_2} \left[ \frac{x_2 f(x_1, x_2)}{f_1(x_1)} \right] dx_2, \qquad \ldots(**)$$

Unconditional mean of $X_2$ is given by

$$E(X_2) = \int_{x_2} x_2 \, f_2(x_2) \, dx_2, \qquad \ldots(***)$$

where $f_2(\cdot)$ is marginal p.d.f. of $X_2$.

From (**) and (***), we conclude that the conditional mean of $X_2$ (given $X_1$) will coincide with unconditional mean of $X_2$ only if

$$\frac{f(x_1, x_2)}{f_1(x_1)} = f_2(x_2)$$

$$\Rightarrow \quad f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$$

*i.e.*, if $X_1$ and $X_2$ are (stochastically) independent.

*(b)* $f(x_1, x_2) = 21\, x_1^2 x_2^3 \,; \; 0 < x_1 < x_2 < 1$
$\qquad = 0, \qquad$ otherwise

Marginal p.d.f. of $X_2$ is given by

$$f_2(x_2) = \int_0^{x_2} f(x_1, x_2)\, dx_1 = 21\, x_2^3 \int_0^{x_2} x_1^2 \, dx_1$$

$$= 21\, x_2^3 \left| \frac{x_1^3}{3} \right|_0^{x_2} = 7\, x_2^6 \,; \quad 0 < x_2 < 1$$

$\therefore$ Conditional p.d.f. of $X_1$ (given $X_2$) is given by

$$f_1(x_1 \mid x_2) = \frac{f(x_1, x_2)}{f_2(x_2)} = 3\,\frac{x_1^2}{x_2^3} \,; \quad 0 < x_1 < x_2 \,; 0 < x_2 < 1$$

Conditional mean of $X_1$ is

$$E(X_1 \mid X_2 = x_2) = \int_0^{x_2} x_1 f_1(x_1 \mid x_2)\, dx_1 = \frac{3}{x_2^3} \int_0^{x_2} x_1^3 \, dx_1$$

$$= \frac{3}{x_2^3} \cdot \left| \frac{x_1^4}{4} \right|_0^{x_2} = \frac{3x_2}{4} \,; \quad 0 < x_2 < 1$$

Now

$$E(X_1^2 \mid X_2 = x_2) = \int_0^{x_2} x_1^2 f_1(x_1 \mid x_2)\, dx_1 = \frac{3}{x_2^3} \int_0^{x_2} x_1^4 \, dx_1$$

$$= \frac{3}{x_2^3} \cdot \frac{x_2^5}{5} = \frac{3}{5} x_2^2$$

$$\therefore \quad \text{Var}(X_1 \mid X_2 = x_2) = E(X_1^2 \mid X_2 = x_2) - [E(X_1 \mid X_2 = x_2)]^2$$

$$= \frac{3}{5} x_2^2 - \frac{9}{16} x_2^2 = \frac{3}{80} x_2^2 \,; \quad 0 < x_2 < 1.$$

**Example 6·35.** *Two random variables X and Y have the followinig joint probability density function :*

$$f(x, y) = 2 - x - y\,;\ 0 \le x \le 1,\ 0 \le y \le 1$$
$$= 0,\ otherwise$$

*Find* (i) *Marginal probability density functions of X and Y.*
(ii) *Conditional density functions.*
(iii) *Var (X) and Var (Y).*
(iv) *Co-variance between X and Y.*

**[Dibrugarh Univ. B.Sc. (Hons.), 1991]**

**Solution.** (i) $f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy$

$$= \int_0^1 (2 - x - y)\, dy = \frac{3}{2} - x$$

$\therefore$ $f_X(x) = \frac{3}{2} - x,\ 0 < x < 1$

$= 0$, otherwise

Similarly $f_Y(y) = \frac{3}{2} - y,\ 0 < y < 1$

$= 0$, otherwise

(ii) $$f_{X|Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{(2 - x - y)}{(3/2 - y)},\ 0 < (x, y) < 1$$

$$f_{Y|X}(y \mid x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{(2 - x - y)}{(3/2 - x)},\ 0 < (x, y) < 1$$

$$E(X) = \int_0^1 x\, f_X(x)\, dx = \int_0^1 x\left(\frac{3}{2} - x\right) dx = \frac{5}{12}$$

$$E(Y) = \int_0^1 y\, f_Y(y)\, dy = \int_0^1 y\left(\frac{3}{2} - y\right) dy = \frac{5}{12}$$

(iii) $$E(X^2) = \int_0^1 x^2\left(\frac{3}{2} - x\right) dx = \left[\frac{3}{6}x^3 - \frac{x^4}{4}\right]_0^1 = \frac{1}{4}$$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{1}{4} - \frac{25}{144} = \frac{11}{144}$$

Similarly $V(Y) = \frac{11}{144}$

(iv) $$E(XY) = \int_0^1 \int_0^1 xy\,(2 - x - y)\, dx\, dy$$

$$= \int_0^1 \left| 2\frac{x^2 y}{2} - \frac{x^3 y}{3} - \frac{x^2 y^2}{2} \right|_0^1 dy$$

$$= \int_0^1 \left(\frac{2}{3}y - \frac{1}{2}y^2\right) dy$$

$$= \left| \frac{y^2}{3} - \frac{y^3}{6} \right|_0^1 = \frac{1}{6}$$

$$= \left| \frac{y^2}{3} - \frac{y^3}{6} \right|_0^1 = \frac{1}{6}$$

$\therefore$ $\quad \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{6} - \frac{5}{12} \cdot \frac{5}{12} = -\frac{1}{144}$.

**Example 6·36.** *Let $f(x, y) = 8xy, 0 < x < y < 1$ ; $f(x, y) = 0$, elsewhere. Find (a) $E(Y | X = x)$, (b) $E(XY | X = x)$, (c) $Var(Y | X = x)$.* **[Calcutta Univ. B.Sc. (Maths Hons.), 1988; Delhi Univ. B.Sc. (Maths Hons.), 1990]**

**Solution.** $f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$

$$= 8x \int_x^1 y\, dy$$

$$= 4x(1 - x^2), 0 < x < 1$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

$$= 8y \int_0^y x\, dx$$

$$= 4y^3, 0 < y < 1$$

$$f_{X|Y}(x | y) = \frac{f(x, y)}{f_Y(y)} = \frac{2x}{y^2},\ f_{Y|X}(y | x) = \frac{2y}{1 - x^2},\ 0 < x < y < 1.$$

(a) $\quad E(Y | X = x) = \int_x^1 y \left( \frac{2y}{1 - x^2} \right) dy = \frac{2}{3} \left( \frac{1 - x^3}{1 - x^2} \right) = \frac{2}{3} \left( \frac{1 + x + x^2}{1 + x} \right)$

(b) $\quad E(XY | X = x) = x\, E(Y | X = x) = \frac{2}{3} \cdot \frac{x(1 + x + x^2)}{(1 + x)}$

(c) $\quad E(Y^2 | X = x) = \int_x^1 y^2 \left( \frac{2y}{1 - x^2} \right) dy = \frac{1}{2} \left( \frac{1 - x^4}{1 - x^2} \right) = \frac{1 + x^2}{2}$

$$\text{Var}(Y | X = x) = E(Y^2 | X = x) - [E(Y | X = x)]^2$$

$$= \frac{1 + x^2}{2} - \frac{4}{9} \cdot \frac{(1 + x + x^2)^2}{(1 + x)^2}$$

## EXERCISE 6(b)

1. The joint probability distribution of $X$ and $Y$ is given by the following table :

| $X \backslash Y$ | 1 | 3 | 9 |
|---|---|---|---|
| 2 | 1/8 | 1/24 | 1/12 |
| 4 | 1/4 | 1/4 | 0 |
| 6 | 1/8 | 1/24 | 1/12 |

(i) Find the marginal probability distribution of $Y$.

*(ii)* Find the conditional distribution of $Y$ given that $X = 2$,

*(iii)* Find the covariance of $X$ and $Y$, and

*(iv)* Are $X$ and $Y$ independent ?

**2.** A fair coin is tossed four times. Let $X$ denote the number of heads occurring and let $Y$ denote the longest string of heads occurring.

*(i)* Determine the joint distribution of $X$ and $Y$, *and* *(ii)* Find Cov $(X, Y)$.

**Hint.**

| $X$ \ $Y$ | 0 | 1 | 2 | 3 | 4 | *Total* |
|---|---|---|---|---|---|---|
| 0 | 1/16 | 0 | 0 | 0 | 0 | 1/16 |
| 1 | 0 | 4/16 | 0 | 0 | 0 | 4/16 |
| 2 | 0 | 3/16 | 3/16 | 0 | 0 | 6/16 |
| 3 | 0 | 0 | 2/16 | 2/16 | 0 | 4/16 |
| 4 | 0 | 0 | 0 | 0 | 1/16 | 1/16 |
| *Total* | 1/16 | 7/16 | 5/16 | 2/16 | 1/16 | 1 |

*(ii)* Cov $(X, Y) = 0{\cdot}85$.

**3.** $X$ and $Y$ are jointly discrete random variables with probability function

$$p(x, y) = 1/4 \text{ at } (x, y) = (-3, -5), (-1, -1), (1, 1), (3, 5)$$
$$= 0, \text{ otherwise}$$

Compute $E(X), E(Y), E(XY)$ and $E(X \mid Y)$. Are $X$ and $Y$ independent ?

**4.** $X_1$ and $X_2$ have a bivariate distribution given by

$$P(X_1 = x_1 \cap X_2 = x_2) = \frac{x_1 + 3x_2}{24}, \text{ where } (x_1, x_2) = (1, 1), (1, 2), (2, 1), (2, 2)$$

Find the conditional mean and variance of $X_1$, given $X_2 = 2$.

**5.** Two random variables $X$ and $Y$ have the following joint probability density function :

$$f(x, y) = k(4 - x - y) \; ; \; 0 \le x \le 2 \; ; \; 0 \le y \le 2$$
$$= 0, \text{ otherwise}$$

Find *(i)* the constant k,

*(ii)* marginal density functions of $X$ and $Y$,

*(iii)* conditional density functions, and

*(iv)* Var $(X)$, Var $(Y)$ and Cov $(X, Y)$. (Poona Univ. B.Sc., Oct. 1991)

**6.** Let the joint probability density function of the random variables $X$ and $Y$ be

$$f(x, y) = 2(x + y - 3xy^2) \; ; \; 0 < x < 1, 0 < y < 1$$
$$= 0, \text{ otherwise}$$

*(i)* Find the marginal distributions of $X$ and $Y$.

*(ii)* Is $E(XY) = E(X)\,E(Y)$ ?

(*iii*) Find $E(X+Y)$ and $E(X-Y)$. **[Calicut Univ. B.Sc., Oct. 1990]**

7. (*a*) Let $X$ and $Y$ have the joint probability density function

$$f(x, y) = 2, 0 < x < y < 1$$
$$= 0, \text{ otherwise}$$

Show that the conditional mean and variance of $X$ given $Y = y$ are $y/2$ and $y^2/12$ respectively.

(*b*) If $f(x, y) = 2$; $0 < x < y, 0 < y < 1$

Find (*i*) $E(Y \mid X)$, (*ii*) $E(X \mid Y)$.

8. Give an example to show that $E(Y)$ may not exist though $E(XY)$ and $E(Y \mid X)$ may both exist? **[Delhi Univ. B.A. (Stat. Hons.) Spl. Course, 1985]**

**Hint.** Consider the joint p.d.f. :

$$f(x, y) = x \cdot e^{-x(1+y)}; x \geq 0, y \geq 0.$$
$$= 0, \text{ otherwise.}$$

Then we shall get :

$$f_X(x) = \int_0^\infty f(x, y)\, dy = e^{-x}; x \geq 0$$

$$f_Y(y) = \int_0^\infty f(x, y)\, dx = \frac{1}{(1+y)^2}; \; y \geq 0$$

$$f(y \mid x) = \frac{f(x, y)}{f_X(x)} = x\, e^{-xy}; \; y \geq 0$$

$$\therefore \quad E(Y) = \int_0^\infty y f(y) = \infty \quad \Rightarrow \quad E(Y) \text{ does not exist.}$$

$$E(XY) = \int_0^\infty \int_0^\infty xy \cdot f(x\, y)\, dx\, dy$$

$$E(Y \mid X = x) = \int_0^\infty y \cdot f(y \mid x)\, dy = \frac{1}{x}$$

$\Rightarrow$ Both $E(XY)$ and $(E(Y \mid X = x)$ exist, though $E(Y)$ does not exist.

9. Three coins are tossed. Let $X$ denote the number of heads on the first two coins, $Y$ denote the number of tails on the last two and $Z$ denote the number of heads on the last two.

(*a*) Find the joint distribution of (*i*) $X$ and $Y$, (*ii*) $X$ and $Z$.

(*b*) Find the conditional distribution of $Y$ given $X = 1$.

(*c*) Find $E(Z \mid X = 1)$.

(*d*) Find $\rho_{X,Y}$ and $\rho_{X,Z}$.

(*e*) Give a joint distribution that is not the joint distribution of $X$ and $Z$ in (*a*) and yet has the same marginals as $f(x, z)$ has in part (*a*).

**[Delhi Univ. B.Sc. (Maths Hons.), 1989]**

**Hint.** The sample space is $S' = \{H, T\} \times \{H, T\} \times \{H, T\}$

$= \{H, T\} \times \{HH, HT, TH, TT\}$

and each of the 8 sample points (outcomes) has the probability $p = 1/8$ of occurrence.

$X$ : Number of heads on the 1st two coins.

$Y$ : Number of tails on the last two coins.

$Z$ : Number of heads on the last two coins.

Then the distribution of $X$, $Y$ and $Z$ is given below :

| *Sample Point :* | *HHH* | *HHT* | *HTH* | *HTT* | *THH* | *THT* | *TTH* | *TTT* |
|---|---|---|---|---|---|---|---|---|
| *Probability* | $p$ | $p$ | $p$ | $p$ | $p$ | $p$ | $p$ | $p$ |
| $X$ | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 0 |
| $Y$ | 0 | 1 | 1 | 2 | 0 | 1 | 1 | 2 |
| $Z$ | 2 | 1 | 1 | 0 | 2 | 1 | 1 | 0 |

Joint Distribution of $X$ and $Y$

| | | $y$ 0 | 1 | 2 | *Total* ($f_x$) |
|---|---|---|---|---|---|
| | 0 | 0 | 1/8 | 1/8 | 1/4 |
| $x$ | 1 | 1/8 | 2/8 | 1/8 | 1/2 |
| | 2 | 1/8 | 1/8 | 0 | 1/4 |
| *Total* ($f_y$) | | 1/4 | 1/2 | 1/4 | 1 |

Joint Distribution of $X$ and $Z$

| | | $z$ 0 | 1 | 2 | *Total* ($f_x$) |
|---|---|---|---|---|---|
| | 0 | 1/8 | 1/8 | 0 | 1/4 |
| $x$ | 1 | 1/8 | 2/8 | 1/8 | 1/2 |
| | 2 | 0 | 1/8 | 1/8 | 1/4 |
| *Total* ($f_z$) | | 1/4 | 1/2 | 1/4 | 1 |

(*b*) $\quad P(Y=0 \mid X=1) = \dfrac{P(Y=0,\ X=1)}{P(X=1)} = \dfrac{1/8}{1/2} = \dfrac{1}{4}$

Similarly, $P(Y=1 \mid X=1) = \dfrac{2/8}{1/2} = \dfrac{1}{2}$ ; $P(Y=2 \mid X=1) = \dfrac{1/8}{1/2} = \dfrac{1}{4}$

(*c*) $\quad E(Z \mid X=1) = \Sigma z \,.\, P(Z \mid X=1) = 0 \times \dfrac{1/8}{1/2} + 1 \times \dfrac{2/8}{1/2} + 2 \times \dfrac{1/8}{1/2} = 1$

(*d*)
$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{-1/4}{\sqrt{1/2 \times 1/2}} = -\frac{1}{2}\,.$$

$$\rho_{XZ} = \frac{\text{Cov}(X, Z)}{\sigma_X \sigma_Z} = \frac{-1/4}{\sqrt{1/2 \times 1/2}} = -\frac{1}{2}$$

(*e*) Let $0 \le \varepsilon \le 1/8$. The joint probability distribution of $(X, Z)$ given belc has the same marginals as in part (*a*).

| | | $Z$ 0 | 1 | 2 | Total ($f_x$) |
|---|---|---|---|---|---|
| $X$ | 0 | 1/8 | 1/8 | 0 | 1/4 |
| | 1 | 1/8 | $2/8 + \varepsilon$ | $1/8 - \varepsilon$ | 1/2 |
| | 2 | 0 | $1/8 - \varepsilon$ | $1/8 + \varepsilon$ | 1/4 |
| Total ($f_z$) | | 1/4 | 1/2 | 1/4 | 1 |

**10.** Let $f_{XY}(x, y) = e^{-(x+y)}$; $0 < x < \infty, 0 < y < \infty$

Find :

(a) $P(X > 1)$ (d) $m$ so that $P(X + Y < m) = ½$

(b) $P(1 < X + Y < 2)$ (e) $P(0 < X < 1 \mid Y = 2)$

(c) $P(X < Y \mid X < 2Y)$ (f) $\rho_{XY}$

**Ans.** $f_X(x) = e^{-x}$; $x \geq 0$ ; $f_Y(y) = e^{-y}$; $y \geq 0$

(a) $1/e$ (b) *Hint.* $X + Y$ is a Gamma variabte with parameter $n = 2$. [See Chapter 8] $(2/e - 3/e^2)$.

(c) $P(X < Y \mid X < 2Y) = \dfrac{P(X < Y \cap X < 2Y)}{P(X < 2Y)} = \dfrac{P(X < Y)}{P(X < 2Y)} = \dfrac{½}{⅔} = \dfrac{3}{4}$

(d) Use hint in (b). $e^{-m}(1 + m) = ½$ ; (e) $(e - 1)/e$

(f) $f_{XY}(x, y) = f_X(x) f_Y(y) \Rightarrow X$ and $Y$ are independent $\Rightarrow \rho_{XY} = 0$.

**11.** The joint p.d.f. of $X$ and $Y$ is given by :

$$f(x, y) = 3(x + y) \;;\; 0 \leq x \leq 1,\; 0 \leq y \leq 1;\; 0 \leq x + y \leq 1$$

Find : (a) Marginal density of $X$. (b) $P(X + Y < ½)$

(c) $E(Y \mid X = x)$ (d) Cov $(X, Y)$

**Ans.** (a) $f_X(x) = \dfrac{3}{2}(1 - x^2)$ ; $0 \leq x \leq 1$.

(b) $P(X + Y < ½) = \displaystyle\int_0^{0\cdot5}\left[\int_0^{0\cdot5-x} 3(x + y)\,dy\right]dx = \frac{1}{8}$

(c) $\dfrac{(1 - x)(x + 2)}{3(1 + x)}$ (d) $E(XY) = \displaystyle\int_0^1\left[\int_0^{1-x} xy\, f(x, y)dy\,dx = \frac{1}{10}\right]$

$$\text{Cov}(x, y) = E(XY) - E(X)E(Y) = \frac{1}{10} - \frac{3}{8} \times \frac{3}{8} = -\frac{13}{320}.$$

**6·10. Moment Generating Function.** The moment generating function (m.g.f.) of a random variable $X$ (about origin) having the probability function $f(x)$ is given by

$$(6\cdot54)\ldots\quad \begin{cases} M_X(t) = E(e^{tX}) = \int e^{tx} f(x)\,dx, & \text{(for } \textit{continuous probability distribution}) \\ \qquad = \sum\limits_x e^{tx} f(x), & \text{(for } \textit{discrete probability distribution}) \end{cases}$$

the integration or summation being extended to the entire range of $x$, $t$ being the real parameter and it is being assumed that the right-hand side of (6·54) is absolutely convergent for some positive number $h$ such that $-h < t < h$. Thus

$$M_X(t) = E(e^{tX}) = E\left[1 + tX + \frac{t^2X^2}{2!} + \ldots + \frac{t^rX^r}{r!} + \ldots\right]$$

$$= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \ldots + \frac{t^r}{r!}E(X^r) + \ldots$$

$$= 1 + t\,\mu_1' + \frac{t^2}{2!}\mu_2' + \dots + \frac{t^r}{r!}\mu_r' + \dots \qquad \dots(6\cdot55)$$

where $\mu_r' = E(X^r) = \int x^r f(x)\,dx$, *for continuous distribution*

$= \sum_x x^r p(x)$, *for discrete distribution*,

is the $r$th moment of $X$ about origin. Thus we see that the coefficient of $\frac{t^r}{r!}$ in $M_X(t)$ gives $\mu_r'$ (above origin). Since $M_X(t)$ generates moments, it is known as moment generating function.

Differentiating (6·55) w.r.t. $t$ and then putting $t = 0$, we get

$$\left[\frac{d^r}{dt^r}\{M_X(t)\}\right]_{t=0} = \left[\frac{\mu_r'}{r!}\cdot r! + \mu'_{r+1}\,t + \mu'_{r+2}\cdot\frac{t^2}{2!} + \dots\right]_{t=0}$$

$$\Rightarrow \qquad \mu_r' = \left[\frac{d^r}{dt^r}\{M_X(t)\}\right]_{t=0} \qquad \dots(6\cdot56)$$

In general, the moment generating function of $X$ about the point $X = a$ is defined as

$$M_X(t)\ (\text{about } X = a) = E\left[e^{t(X-a)}\right]$$

$$= E\left[1 + t(X-a) + \frac{t^2}{2!}(X-a)^2 + \dots + \frac{t^r}{r!}(X-a)^r + \dots\right]$$

$$= 1 + t\,\mu_1' + \frac{t^2}{2!}\mu_2' + \dots + \frac{t^r}{r!}\mu_r' + \dots \qquad \dots(6\cdot57)$$

where $\mu_r' = E\{(X-a)^r\}$, is the $r$th moment about the point $X = a$.

**6·10·1. Some Limitations of Moment Generating Functions.** Moment generating function suffers from some drawbacks which have restricted its use in Statistics. We give below some of the deficiencies of m.g.f.'s with illustrative examples.

1. *A random variable X may have no moments although its m.g.f. exists.* For example, let us consider a discrete r.v. with probability function

$$f(x) = \begin{cases} \dfrac{1}{x(x+1)}\ ; & x = 1, 2, \dots \\ 0, & \text{otherwise} \end{cases}$$

Here
$$E(X) = \sum_{x=1}^{\infty} x f(x) = \sum_{x=1}^{\infty} \frac{1}{(x+1)}$$

$$= \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots$$

$$= \left[\sum_{x=1}^{\infty} \frac{1}{x}\right] - 1$$

Since $\sum_{x=1}^{\infty} \frac{1}{x}$ is a divergent series, $E(X)$ does not exist and consequently no moment of $X$ exists. However, the m.g.f. of $X$ is given by

$$
\begin{aligned}
M_X(t) &= \sum_{x=1}^{\infty} e^{tx} \cdot f(x) = \sum_{x=1}^{\infty} \frac{e^{tx}}{x(x+1)} \\
&= \sum_{x=1}^{\infty} \frac{z^x}{x(x+1)}, \quad (z = e^t) \qquad \ldots(*) \\
&= \frac{z}{1.2} + \frac{z^2}{2.3} + \frac{z^3}{3.4} + \frac{z^4}{4.5} + \ldots \\
&= z\left[1 - \frac{1}{2}\right] + z^2\left[\frac{1}{2} - \frac{1}{3}\right] + z^3\left[\frac{1}{3} - \frac{1}{4}\right] + z^4\left[\frac{1}{4} - \frac{1}{5}\right] + \ldots \\
&= \left[z + \frac{z^2}{2} + \frac{z^3}{3} + \frac{z^4}{4} + \ldots\right] - \frac{z}{2} - \frac{z^2}{3} - \frac{z^3}{4} - \frac{z^4}{5} - \ldots \\
&= -\log(1-z) - \frac{1}{z}\left[\frac{z^2}{2} + \frac{z^3}{3} + \frac{z^4}{4} + \ldots\right] \\
&= -\log(1-z) + 1 + \frac{1}{z}\log(1-z), \quad |z| < 1 \\
&= 1 + \left[\frac{1}{z} - 1\right]\log(1-z), \quad |z| < 1 \\
&= 1 + (e^{-t} - 1)\log(1 - e^t), \quad t < 0
\end{aligned}
$$

$$[\because |z| < 1 \Rightarrow |e^t| < 1 \Rightarrow t < 0]$$

And $\quad M_X(t) = 1$, for $t = 0$, [From (*)]

while for $t > 0$, $M_X(t)$ does not exist.

**2.** *A random variable X can have m.g.f. and some (or all) moments, yet the m.g.f. does not generate the moments.* For example, consider a discrete r.v. with probability function

$$P(X = 2^x) = \frac{e^{-1}}{x!}\ ; \ x = 0, 1, 2, \ldots$$

Here $$E(X^r) = \sum_{\xi=0}^{\infty} (2^x)^r P(X = 2^x) = e^{-1} \sum_{\xi=0}^{\infty} \frac{(2^r)^x}{x!}$$

$$= e^{-1} \cdot \exp(2^r) = \exp(2^r - 1)$$

Hence all the moments of $X$ exist.

The m.g.f. of $X$, if it exists is given by

$$M_X(t) = \sum_{x=0}^{\infty} \exp(t \cdot 2^x)\left(\frac{e^{-1}}{x!}\right) = e^{-1} \sum_{x=0}^{\infty} \exp(t \cdot 2^x)\frac{1}{x!}$$

By D'Alembert's ratio test, the series on the R.H.S. converges for $t \le 0$ and diverges for $t > 0$. Hence $M_X(t)$ cannot be differentiated at $t = 0$ and has no Maclaurin's expansion and consequently it does not generate moments.

**3.** *A r.v. X can have all or some moments; but m.g.f. does not exist except perhaps at one point.*

For example, let $X$ be a r.v. with probability function

$$P(X = \pm 2^x) = \frac{e^{-1}}{2x!}\ ; \quad x = 0, 1, 2, \ldots$$
$$= 0, \text{ otherwise.}$$

Since the distribution is symmetric about the line $X = 0$, all moments of odd order about origin vanish, *i.e.*,

$$E(X^{2r+1}) = 0 \quad \Rightarrow \quad \mu_{2r+1} = 0$$

$$E(X^{2r}) = \sum_{x=0}^{\infty} (\pm 2^x)^{2r} \left(\frac{1}{2ex!}\right) = \frac{1}{e} \sum_{x=0}^{\infty} \frac{(2^{2r})^x}{x!}$$
$$= \frac{1}{e} \cdot \exp(2^{2r}) = \exp(2^{2r} - 1)$$

Thus all the moments of $X$ exist. The m.g.f. of $X$, if it exists, is given by

$$M_X(t) = \sum_{x=0}^{\infty} \left[ (e^{t.2^x} + e^{-t.2^x}) \frac{1}{2e\,x!} \right]$$
$$= e^{-1} \sum_{x=0}^{\infty} \left[ \frac{\cos(t.2^x)}{x!} \right]$$

which converges only for $t = 0$.

As an illustration of a continuous probability distribution, consider Pareto distribution with p.d.f.

$$p(x) = \frac{\theta \cdot a^\theta}{x^{\theta+1}}\ ; \qquad x \geq a\ ;\ \theta > 1$$

$$E(X^r) = \theta \cdot a^\theta \int_a^{\infty} x^{r-\theta-1}\, dx = \theta \cdot a^\theta \cdot \left| \frac{x^{r-\theta}}{r-\theta} \right|_a^{\infty},$$

which is finite iff $r - \theta < 0 \quad \Rightarrow \quad \theta > r$ and then

$$E(X^r) = \theta\, a^\theta \left[ 0 - \frac{a^{r-\theta}}{r-\theta} \right] = \frac{\theta \cdot a^r}{\theta - r}\ ; \quad \theta > r$$

However, the m.g.f. is given by :

$$M_X(t) = \theta \cdot a^\theta \int_a^{\infty} \frac{e^{tx}}{x^{\theta+1}}\, dx,$$

which does not exist, since $e^{tx}$ dominates $x^{\theta+1}$ and $(e^{tx}/x^{\theta+1}) \to \infty$ as $x \to \infty$ and hence the integral is not convergent.

For more illustrations see Student's t-distribution and Snedecor's F-distributions, for which m.g.f.'s do not exist, though the moments of all orders exist. [c.f. Chapter 14, § 14·2·4 and 14·5·2.].

**Remark.** The reason that m.g.f. is a poor tool in comparison with characteristic function (c.f. § 6·12) is that the domain of the dummy parameter '$t$' of the m.g.f. depends on the distribution of the r.v. under consideration, while characteristic function exists for all real $t$, $(-\infty < t < \infty)$. If m.g.f. is valid for $t$ lying in an interval containing zero, then m.g.f. can be expanded with perhaps some additional restrictions.

**6·10·2. Theorems on Moment Generating Functions.**

**Theorem 6·17.** $M_{cX}(t) = M_X(ct)$, $c$ being a constant. ...(6·58)

**Proof.** By def.,

$$\text{L.H.S.} = M_{cX}(t) = E(e^{t.cX})$$
$$\text{R.H.S.} = M_X(ct) = E(e^{ctX}) = \text{L.H.S.}$$

**Theorem 6·18.** *The moment generating function of the sum of a number of independent random variables is equal to the product of their respective moment generating functions.*

Symbolically, if $X_1, X_2, ..., X_n$ are independent random variables, then the moment generating function of their sum $X_1 + X_2 + ... + X_n$ is given by

$$M_{X_1+X_2+...+X_n}(t) = M_{X_1}(t)\, M_{X_2}(t) \dots M_{X_n}(t) \quad ...(6·59)$$

**Proof.** By definition,

$$M_{X_1+X_2+...+X_n}(t) = E\left[e^{t(X_1+X_2+...+X_n)}\right]$$
$$= E\left[e^{tX_1} . e^{tX_2} \dots e^{tX_n}\right]$$
$$= E(e^{tX_1})\, E(e^{tX_2}) \dots E(e^{tX_n})$$
$$(\because X_1, X_2, ..., X_n \text{ are independent})$$
$$= M_{X_1}(t) . M_{X_2}(t) \dots M_{X_n}(t)$$

Hence the theorem.

**Theorem 6·19.** *Effect of change of origin and scale on M.G.F.* Let us transform $X$ to the new variable $U$ by changing both the origin and scale in $X$ as follows :

$$U = \frac{X - a}{h}, \text{ where } a \text{ and } h \text{ are constants}$$

M.G.F. of $U$ (about origin) is given by

$$M_U(t) = E(e^{tU}) = E\left[\exp\{t(x-a)/h\}\right]$$
$$= E\left[e^{tX/h} . e^{-at/h}\right] = e^{-at/h} E(e^{tX/h})$$
$$= e^{-at/h} E(e^{Xt/h}) = e^{-at/h} M_X(t/h) \quad ...(6·60)$$

where $M_X(t)$ is the m.g.f. of $X$ about origin.

In particular, if we take $a = E(X) = \mu$ (say) and $h = \sigma_X = \sigma$ (say), then

$$U = \frac{X - E(X)}{\sigma_X} = \frac{X - \mu}{\sigma} = Z \text{ (say)},$$

is known as a *standard variate.* Thus the m.g.f. of a standard variate $Z$ is given by

$$M_Z(t) = e^{-\mu t/\sigma} M_X(t/\sigma) \quad ...(6·61)$$

**Remark.** $E(Z) = E\left(\dfrac{X - \mu}{\sigma}\right) = \dfrac{1}{\sigma} E(X - \mu)$

$$= \frac{1}{\sigma}\{E(X) - \mu\} = \frac{1}{\sigma}(\mu - \mu) = 0$$

and $\quad V(Z) = V\left(\dfrac{X-\mu}{\sigma}\right) = \dfrac{1}{\sigma^2} V(X-\mu)$ [c.f. Cor. (i) Theorem 6·8]

$= \dfrac{1}{\sigma^2} V(X) = \dfrac{1}{\sigma} \sigma^2 = 1$ [c.f. Cor. (iii) Theorem 6·8]

$\therefore \quad E(Z) = 0$ and $V(Z) = 1$, *i.e.*, the mean and variance of a standard variate are 0 and 1 respectively.

**6·10·3. Uniqueness Theorem of Moment Generating Function.** *The moment generating function of a distribution, if it exists, uniquelly determines the distribution.* This implies that corresponding to a given probability distribution, there is only one m.g.f. (provided it exists) and corresponding to a given m.g.f., there is only one probability distribution. Hence $M_X(t) = M_Y(t) \Rightarrow X$ and $Y$ are identically distributed. [For detailed discussion, see Uniqueness Theorem of Characteristic Functions – Theorem 6·27, page 6·90]

**6·11. Cumulants.** Cumulants generating function $K(t)$ is defined as

$$K_X(t) = \log_e M_X(t), \qquad \text{...(6·62)}$$

provided the right-hand side can be expanded as a convergent series in powers of $t$. Thus

$$K_X(t) = \kappa_1 t + \kappa_2 \frac{t^2}{2!} + \dots + \kappa_r \frac{t^r}{r!} + \dots = \log M_X(t)$$

$$= \log \left[ 1 + \mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \dots + \mu_r' \frac{t^r}{r!} + \dots \right] \qquad \text{...(6·62a)}$$

where $\kappa_r =$ coefficient of $\dfrac{t^r}{r!}$ in $K_X(t)$ is called the $r$th cumulant. Hence

$$\kappa_1 t + \kappa_2 \frac{t^2}{2!} + \kappa_3 \frac{t^3}{3!} + \kappa_4 \frac{t^4}{4!} + \dots$$

$$= \left[ \left( \mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \mu_4' \frac{t^4}{4!} + \dots \right) \right.$$

$$- \frac{1}{2} \left( \mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \dots \right)^2$$

$$\left. + \frac{1}{3} \left( \mu_1' t + \mu_2' \frac{t^2}{2!} + \dots \right)^3 - \frac{1}{4} \left( \mu_1' t + \mu_2' \frac{t^2}{2!} + \dots \right)^4 + \dots \right]$$

Comparing the coefficients of like powers of '$t$' on both sides, we get the relationship between the moments and cumulants. Hence, we have

$$\kappa_1 = \mu_1' = \text{Mean}, \quad \frac{\kappa_2}{2!} = \frac{\mu_2'}{2!} - \frac{\mu_1'^2}{2!} \quad \Rightarrow \quad \kappa_2 = \mu_2' - \mu_1'^2 = \mu_2$$

$$\frac{\kappa_3}{3!} = \frac{\mu_3'}{3!} - \frac{1}{2} \frac{2\mu_1' \mu_2'}{2!} + \frac{\mu_1'^3}{3} \quad \Rightarrow \quad \kappa_3 = \mu_3' - 3\mu_2' \mu_1' + 2\mu_1'^3 = \mu_3$$

Also

$$\frac{\kappa_4}{4!}=\frac{\mu_4'}{4}-\frac{1}{2}\left(\frac{\mu_2'^2}{4}+\frac{2\mu_1'\mu_3'}{3!}\right)+\frac{1}{3}\frac{3\mu_1'^2\mu_2'}{2}-\frac{\mu_1'^4}{4}$$

$$\Rightarrow \kappa_4=\mu_4'-3\mu_2'^2-4\mu_1'\mu_3'+12\mu_1'^2\mu_2'-6\mu_1'^4$$

$$=(\mu_4'-4\mu_3'\mu_1'+6\mu_2'\mu_1'^2-3\mu_1'^4)-3(\mu_2'^2-2\mu_2'\mu_1'^2+\mu_1'^4)$$

$$=\mu_4-3(\mu_2'-\mu_1'^2)^2=\mu_4-3\mu_2^2=\mu_4-3\kappa_2^2 \qquad (\because \mu_2=\kappa_2)$$

$$\Rightarrow \qquad \mu_4=\kappa_4+3\kappa_2^2$$

Hence we have obtained:

$$\left.\begin{aligned} &\text{Mean}=\kappa_1\\ &\mu_2=\kappa_2=\text{Variance}\\ &\mu_3=\kappa_3\\ &\mu_4=\kappa_4+3\kappa_2^2\end{aligned}\right\} \qquad \text{...(6·62}b\text{)}$$

**Remarks.** 1. These results are of fundamental importance and should be committed to memory.

2. If we differentiate both sides of (6·62$a$) w.r.t. $t$ '$r$' times and then put $t=0$, we get

$$\kappa_r=\left[\frac{d^r}{dt^r}K_X(t)\right]_{t=0} \qquad \text{...(6·62}c\text{)}$$

**6·11·1. Additive Property of Cumulants.** *The rth cumulant of the sum of independent random variables is equal to the sum of the rth cumulants of the individual variables.* Symbolically,

$$\kappa_r(X_1+X_2+\ldots+X_n)=\kappa_r(X_1)+\kappa_r(X_2)+\ldots+\kappa_r(X_n), \qquad \text{...(6·63)}$$

where $X_i$; $i=1,2,\ldots,n$ are independent random variables.

**Proof.** We have, since $X_i$'s are independent.

$$M_{X_1+X_2+\ldots+X_n}(t)=M_{X_1}(t)\,M_{X_2}(t)\ldots M_{X_n}(t)$$

Taking logarithm of both sides, we get

$$K_{X_1+X_2+\ldots+X_n}(t)=K_{X_1}(t)+K_{X_2}(t)+\ldots+K_{X_n}(t)$$

Differentiating both sides w.r.t. $t$ '$r$' times and putting $t=0$, we get

$$\left[\frac{d^r}{dt^r}K_{X_1+X_2+\ldots+X_n}(t)\right]_{t=0}=\left[\frac{d^r}{dt^r}K_{X_1}(t)\right]_{t=0}$$

$$+\left[\frac{d^r}{dt^r}K_{X_2}(t)\right]_{t=0}+\ldots+\left[\frac{d^r}{dt^r}K_{X_n}(t)\right]_{t=0}$$

$$\Rightarrow \qquad \kappa_r(X_1+X_2+\ldots+X_n)=\kappa_r(X_1)+\kappa_r(X_2)+\ldots+\kappa_r(X_n),$$

which establishes the result.

**6·11·2. Effect of Change of Origin and Scale on Cumulants.** If we take

$$U=\frac{X-a}{h}, \text{ then } M_U(t)=\exp(-at/h)M_X(t/h)$$

$$\therefore \quad K_U(t) = -\frac{at}{h} + K_X(t/h)$$

$$\kappa_1' + \kappa_1' t + \kappa_2' \frac{t^2}{2!} + \ldots + \kappa_r' \frac{t^r}{r!} + \ldots = -\frac{at}{h} + \kappa_1 (t/h) + \kappa_2 \frac{(t/h)^2}{2!} + \ldots + \kappa_r \frac{(t/h)^r}{r!} + \ldots$$

where $\kappa_r'$ and $\kappa_r$ are the $r$th cumulants of $U$ and $X$ respectively.

Comparing coefficients, we get

$$\kappa_1' = \frac{\kappa_1 - a}{h} \quad \text{and} \quad \kappa_r' = \frac{\kappa_r}{h^r} \; ; \; r = 2, 3, \ldots \qquad \ldots(6·63a)$$

Thus we see that except the first cumulant, all cumulants are independent of change of origin. But the cumulants are not invariant of change of scale as the $r$th cumulant of $U$ is $(1/h^r)$ times the $r$th cumulant of the distribiution of $X$.

**Example 6·37.** *Let the random variable X assume the value 'r' with the probability law :*

$$P(X = r) = q^{r-1} p \; ; \; r = 1, 2, 3, \ldots$$

*Find the m.g.f. of X and hence its mean and variance.*

[Calicut Univ. B.Sc., Oct. 1992]

**Solution.** $M_X(t) = E(e^{tX})$

$$= \sum_{r=1}^{\infty} e^{tr} \; q^{r-1} p = \frac{p}{q} \sum_{r=1}^{\infty} (qe^t)^r$$

$$= \frac{p}{q} qe^t \sum_{r=1}^{\infty} (qe^t)^{r-1} = pe^t \left[ 1 + qe^t + (qe^t)^2 + \ldots \right]$$

$$= \left( \frac{pe^t}{1 - qe^t} \right)$$

If dash (′) denotes the differentiation w.r.t. $t$, then we have

$$M_X'(t) = \frac{pe^t}{(1 - qe^t)^2}, \quad M_X''(t) = pe^t \frac{(1 + qe^t)}{(1 - qe^t)^3}$$

$$\therefore \quad \mu_1' \text{ (about origin)} = M_X'(0) = \frac{p}{(1-q)^2} = \frac{1}{p}$$

$$\mu_2' \text{ (about origin)} = M_X''(0) = \frac{p(1+q)}{(1-q)^3} = \frac{1+q}{p^2}.$$

Hence $\qquad \text{mean} = \mu_1' \text{ (about origin)} = \dfrac{1}{p}$

and $\qquad \text{variance} = \mu_2 = \mu_2' - \mu_1'^2 = \dfrac{1+q}{p^2} - \dfrac{1}{p^2} = \dfrac{q}{p^2}$

**Example 6·38.** *The probability density function of the random variable X follows the following probability law :*

$$p(x) = \frac{1}{2\theta}\, exp\left(-\frac{|x-\theta|}{\theta}\right), \ -\infty < x < \infty$$

*Find M.G.F. of X. Hence or otherwise find E (X) and V (X).*

**[Punjab Univ. M.A.(Eco.), 1991]**

**Solution.** The moment generating function of $X$ is

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} \frac{1}{2\theta} \exp\left(-\frac{|x-\theta|}{\theta}\right) e^{tx}\, dx$$

$$= \int_{-\infty}^{\theta} \frac{1}{2\theta} \exp\left(-\frac{|\theta-x|}{\theta}\right) e^{tx}\, dx$$

$$+ \int_{\theta}^{\infty} \frac{1}{2\theta} \exp\left(-\frac{|x-\theta|}{\theta}\right) e^{tx}\, dx$$

For $x \in (-\infty, \theta)$, $x-\theta < 0 \Rightarrow \theta - x > 0$

$\therefore \quad |x-\theta| = \theta - x \ \ \forall\, x \in (-\infty, \infty)$

Similarly, $\quad |x-\theta| = x-\theta \ \ \forall\, x \in (\theta, \infty)$

$$\therefore \quad M_X(t) = \frac{e^{-1}}{2\theta} \int_{-\infty}^{\theta} \exp\left[x\left(t+\frac{1}{\theta}\right)\right] dx + \frac{e}{2\theta} \int_{\theta}^{\infty} \exp\left[-x\left(\frac{1}{\theta}-t\right)\right] dx$$

$$= \frac{e^{-1}}{2\theta} \cdot \frac{1}{\left(t+\frac{1}{\theta}\right)} \cdot \exp\left[\theta\left(t+\frac{1}{\theta}\right)\right]$$

$$+ \frac{e}{2\theta} \cdot \frac{1}{\left(\frac{1}{\theta}-t\right)} \cdot \exp\left[-\theta\left(\frac{1}{\theta}-t\right)\right]$$

$$= \frac{e^{\theta t}}{2(\theta t+1)} + \frac{e^{\theta t}}{2(1-\theta t)} = \frac{e^{\theta t}}{1-\theta^2 t^2}$$

$$= e^{\theta t}(1-\theta^2 t^2)^{-1}$$

$$= \left[1 + \theta t + \frac{\theta^2 t^2}{2!} + \ldots\right]\left[1 + \theta^2 t^2 + \theta^4 t^4 + \ldots\right]$$

$$= 1 + \theta t + \frac{3\,\theta^2 t^2}{2!} + \ldots$$

$\therefore \quad E(X) = \mu' = \text{Coefficient of } t \text{ in } M_X(t) = \theta$

$$\mu_2' = \text{Coefficient of } \frac{t^2}{2!} \text{ in } M_X(t) = 3\,\theta^2$$

Hence $\quad \text{Var}(X) = \mu_2' - \mu_1'^2 = 3\,\theta^2 - \theta^2 = 2\,\theta^2$

**Example 6·39.** *If the moments of a variate X are defined by*

$$E(X^r) = 0{\cdot}6\,;\; r = 1, 2, 3, \ldots$$

*show that* $P(X = 0) = 0{\cdot}4, P(X = 1) = 0{\cdot}6, P(X \geq 2) = 0.$

**[Delhi Univ. B.Sc. (Maths Hons.), 1985]**

**Solution.** The m.g.f. of variate $X$ is :

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu_r' = 1 + \sum_{r=1}^{\infty} \frac{t^r}{r!}(0{\cdot}6)$$

$$= 0{\cdot}4 + 0{\cdot}6 \sum_{r=0}^{\infty} \frac{t^r}{r!} = 0{\cdot}4 + 0{\cdot}6\, e^t \qquad \ldots(i)$$

But $\quad M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} P(X = x)$

$$= P(X = 0) + e^t \,.\, P(X = 1) + \sum_{x=2}^{\infty} e^{tx} \,.\, P(X = x) \qquad \ldots(ii)$$

From *(i)* and *(ii)*, we get :

$$P(X = 0) = 0{\cdot}4\;;\; P(X = 1) = 0{\cdot}6\;;\; P(X \geq 2) = 0.$$

**Remark.** In fact *(i)* is the m.g.f. of Bernoulli variate $X$ with $P(X = 0) = q = 0{\cdot}4$ and $P(X = 1) = p = 0{\cdot}6$ [See § 7·1·2] and $P(X \geq 2) = 0$.

**Example 6·40.** *Find the moment generating function of the random variable whose moments are*

$$\mu_r' = (r + 1)!\, 2^r$$

**Solution.** The m.g.f. is given by

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu'_r = \sum_{r=0}^{\infty} \frac{t^r}{r!}(r + 1)!\, 2^r$$

$$= \sum_{r=0}^{\infty} (r + 1)(2t)^r$$

$$M_X(t) = 1 + 2\,.\,(2t) + 3\,(2t)^2 + 4\,(2t)^3 + \ldots$$

$$= (1 - 2t)^{-2}$$

**Aliter.** The R.H.S. is an arithmetic-geometric series with ratio $r = (2t)$

Let $\qquad S = 1 + 2r + 3r^2 + 4r^3 + \ldots$

Then $\qquad rS = \quad r + 2r^2 + 3r^3 + \ldots$

$\therefore \qquad (1 - r)S = 1 + r + r^2 + \ldots = \dfrac{1}{(1 - r)}$

$\Rightarrow \qquad S = \dfrac{1}{(1 - r)^2} = (1 - r)^{-2} = (1 - 2t)^{-2}$

**Remark.** This is the m.g.f. of Chi-square ($\chi^2$) distribution with parameter (degrees of freedom) $n = 2$ [c.f. Chapter 13].

**Example 6·41.** *If $\mu_r'$ is the rth moment about origin, prove that*

$$\mu_r' = \sum_{j=1}^{r} \binom{r-1}{j-1} \mu_{r-j}' \kappa_j,$$

*where $\kappa_j$ is the jth cumulant.*

**Solution.** Differentiating both sides of (6·62*a*) in § 6·11, page 6·72 w.r.t. $t$, we get

$$\kappa_1 + \kappa_2 t + \kappa_3 \frac{t^2}{2\,!} + \ldots + \kappa_r \frac{t^{r-1}}{(r-1)\,!} + \ldots$$

$$= \frac{\mu_1' + \mu_2' t + \mu_3' \dfrac{t^2}{2\,!} + \ldots + \mu_r' \dfrac{t^{r-1}}{(r-1)\,!} + \ldots}{1 + \mu_1' t + \mu_2' \dfrac{t^2}{2\,!} + \ldots + \mu_r' \dfrac{t^r}{r\,!} + \ldots}$$

$$\Rightarrow \quad \left[\kappa_1 + \kappa_2 t + \kappa_3 \frac{t^2}{2\,!} + \ldots + \kappa_r \frac{t^{r-1}}{(r-1)\,!} + \ldots\right]$$

$$\times \left[1 + \mu_1' t + \mu_2' \frac{t^2}{2\,!} + \ldots + \mu_r' \frac{t^r}{r\,!} + \ldots\right]$$

$$= \mu_1' + \mu_2' t + \mu_3' \frac{t^2}{2\,!} + \ldots + \mu_r' \frac{t^{r-1}}{(r-1)\,!} + \ldots$$

Comparing the coefficient of $\dfrac{t^{r-1}}{(r-1)\,!}$ on both sides, we get

$$\mu_r' = \kappa_1 \,.\, \mu_{r-1}' + (r-1)\, \kappa_2 \,.\, \mu_{r-2}' + \binom{r-1}{2} \kappa_3 \,.\, \mu_{r-3}' + \ldots + \kappa_r$$

$$= \binom{r-1}{0} \mu_{r-1}' \kappa_1 + \binom{r-1}{1} \mu_{r-2}' \kappa_2 + \binom{r-1}{2} \mu_{r-3}' \kappa_3$$

$$+ \ldots + \binom{r-1}{r-1} \mu_0' \kappa_r$$

$$= \sum_{j=1}^{r} \binom{r-1}{j-1} \mu_{r-j}' \kappa_j,$$

which is the required result.

**6·12. Characteristic Function.** In some cases m.g.f. does not exist, since the integral $\int_{-\infty}^{\infty} e^{tx} f(x)\, dx$ or the series $\sum_x e^{tx} p(x)$ does not converge absolutely for real values of $t$ for some distributions. For example, for the continuous probability distribution

$$dF(x) = C \frac{1}{(1 + x^2)^m}\, dx \;;\; m > 1, -\infty < x < \infty, \;.$$

the m.g.f. does not exist, since the integral

$$M_X(t) = C \int_{-\infty}^{\infty} e^{tx} \frac{1}{(1 + x^2)^m}\, dx,$$

does not converge absolutely for finite positive values of $m$ because the function $e^{x}$ dominates the function $x^{2m}$ so that $e^{x}/x^{2m} \to \infty$ as $x \to \infty$.

Again, for the discrete probability distribution

$$f(x) = \left.\begin{array}{l} \dfrac{6}{\pi^2 x^2}\,;\ x = 1, 2, 3, \ldots \\ = 0, \text{ elsewhere} \end{array}\right\}$$

$$M_X(t) = \sum_x e^{tx} f(x) = \frac{6}{\pi^2} \sum_{x=1}^{\infty} \left( \frac{e^{tx}}{x^2} \right)$$

The series is not convergent (by D'Alembert's Ratio Test) for $t > 0$. Thus there does not exist a positive number $h$ such that $M_X(t)$ exists for $-h < t < h$. Hence $M_X(t)$ does not exist in this case also.

A more serviceable function than the m.g.f. is what is known as characteristic function and is defined as

$$\phi_X(t) = E(e^{itX}) = \int e^{itx} f(x)\, dx \qquad \textit{(for continuous probability distributions)}$$

$$= \sum_x e^{itx} f(x) \qquad \textit{(for discrete probability distributions)} \qquad \ldots(6{\cdot}64)$$

If $F_X(x)$ is the distsribution function of a continuous random variable $X$, then

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x) \qquad \ldots(6{\cdot}64a)$$

Obviously $\phi(t)$ is a complex valued function of real variable $t$. It may be noted that

$$|\phi(t)| = \left| \int e^{itx} f(x)\, dx \right| \le \int |e^{itx}|\, f(x)\, dx = \int f(x)\, dx = 1,$$

since $\quad |e^{itx}| = |\cos tx + i \sin tx|^{1/2} = (\cos^2 tx + \sin^2 tx)^{1/2} = 1$

Since $|\phi(t)| \le 1$, characteristic function $\phi_X(t)$ always exists.

Yet another advantage of characteristic function lies in the fact that it uniquely determines the distribution function, *i.e.*, if the characteristic function of a distribution is given, the distribution can be uniquely determined by the theorem, known as the Uniqueness Theorem of Characteristic Functions [c.f. Theorem 6·27 page 6·90].

**6·12·1. Properties of Characteristic Functions.** For all real '$t$', we have

*(i)* $\phi(0) = \int_{-\infty}^{\infty} dF(x) = 1 \qquad \ldots(6{\cdot}64b)$

*(ii)* $|\phi(t)| \le 1 = \phi(0) \qquad \ldots(6{\cdot}64c)$

*(iii)* $\phi(t)$ is continuous everywhere, *i.e.*, $\phi(t)$ is a continuous function of $t$ in $(-\infty, \infty)$. Rather $\phi(t)$ is uniformly continuous in '$t$'

**Proof.** For $h \ne 0$, $|\phi_X(t+h) - \phi_X(t)| = \left| \int_{-\infty}^{\infty} \left[ e^{i(t+h)x} - e^{itx} \right] dF(x)\right.$

$$\leq \int_{-\infty}^{\infty} \left| e^{itx} (e^{ihx} - 1) \right| dF(x)$$

$$= \int_{-\infty}^{\infty} \left| e^{ihx} - 1 \right| dF(x) \qquad ...(*)$$

The last integral does not depend on $t$. If it tends to zero as $h \to 0$ then $\phi_X(t)$ is uniformly continuous in '$t$'.

Now $\quad \left| e^{ihx} - 1 \right| \leq \left| e^{ihx} \right| + 1 = 2$

$$\therefore \quad \int_{-\infty}^{\infty} \left| e^{ihx} - 1 \right| dF(x) \leq 2 \int_{-\infty}^{\infty} dF(x) = 2.$$

Hence by Dominated Convergence Theorem (D.C.T.), taking the limit inside the integral sign in (*), we get

$$\lim_{h \to 0} \left| \phi_X(t+h) - \phi_X(t) \right| \leq \int_{-\infty}^{\infty} \lim_{h \to 0} \left| e^{ihx} - 1 \right| dF(x) = 0$$

$$\Rightarrow \quad \lim_{h \to 0} \phi_X(t+h) = \phi_X(t) \; \forall \; t.$$

Hence $\phi_X(t)$ is uniformly continuous in '$t$'.

(*iv*) $\phi_X(-t)$ and $\phi_X(t)$ are conjugate functions, *i.e.*, $\phi_X(-t) = \overline{\phi_X(t)}$, where $\bar{a}$ is the complex conjugate of $a$.

**Proof.** $\qquad \phi_X(t) = E(e^{itX}) = E[\cos tX + i \sin tX]$

$$\Rightarrow \quad \overline{\phi_X(t)} = E[\cos tX - i \sin tX]$$
$$= E[\cos(-t)X + i \sin(-t)X]$$
$$= E(e^{-itX}) = \phi_X(-t).$$

### 6·12·2. Theorems on Characteristic Function.

**Theorem 6·20.** *If the distribution function of a r.v. X is symmetrical about zero, i.e.,*

$$1 - F(x) = F(-x) \quad \Rightarrow \quad f(-x) = f(x),$$

*then $\phi_X(t)$ is real valued and even function of t.*

**Proof.** By definition we have

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx = \int_{-\infty}^{\infty} e^{-ity} f(-y)\, dy \qquad (x = -y)$$

$$= \int_{-\infty}^{\infty} e^{-ity} f(y)\, dy \qquad [\because f(-y) = f(y)]$$

$$= \phi_X(-t) \qquad ...(*)$$

$\Rightarrow \quad \phi_X(t)$ is an even function of $t$.

Also $\qquad \overline{\phi_X(t)} = \phi_X(-t) \qquad$ [c.f. Property (*iv*) § 6·12·1]

$\therefore \qquad \overline{\phi_X(t)} = \phi_X(-t) = \phi_X(t) \qquad$ (From *)

Hence $\phi_X(t)$ is a real valued function of $t$.

**Theorem 6·21.** *If X is some random variable with characteristic function $\phi_X(t)$, and if $\mu_r' = E(X^r)$ exists, then*

$$\mu_r' = (-i)^r \left[\frac{\partial^r}{\partial t^r}\phi(t)\right]_{t=0} \qquad ...(6·65)$$

**Proof.** $\phi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\,dx$

Differentiating (under the integral sign) '$r$' times w.r.t. $t$, we get

$$\frac{\partial^r}{\partial t^r}\phi(t) = \int_{-\infty}^{\infty} (ix)^r \cdot e^{itx} f(x)\,dx = (i)^r \int_{-\infty}^{\infty} x^r e^{itx} f(x)\,dx$$

$$\therefore \left[\frac{\partial^r}{\partial t^r}\phi(t)\right]_{t=0} = (i)^r \left[\int_{-\infty}^{\infty} x^r e^{itx} f(x)\,dx\right]_{t=0}$$

$$= (i)^r \int_{-\infty}^{\infty} x^r f(x)\,dx = i^r E(X^r) = i^r \mu_r'$$

Hence $\mu_r' = \left(\frac{1}{i}\right)^r \left[\frac{\partial^r}{\partial t^r}\phi(t)\right]_{t=0} = (-i)^r \left[\frac{\partial^r}{\partial t^r}\phi(t)\right]_{t=0}$

The theorems, *viz.*, 6·17, 6·18 and 6·19 on m.g.f. can be easily extended to the characteristic functions as given below.

**Theorem 6·22.** $\phi_{cX}(t) = \phi_X(ct)$, *c, being a constant.*

**Theorem 6·23.** *If $X_1$ and $X_2$ are independent random variables, then*

$$\phi_{X_1+X_2}(t) = \phi_{X_1}(t)\ \phi_{X_2}(t) \qquad ...(*)$$

More, generally for independent random variables $X_i$ ; $i = 1, 2, ..., n$, we have

$$\phi_{X_1+X_2+...+X_n}(t) = \phi_{X_1}(t)\ \phi_{X_2}(t) \ldots \phi_{X_n}(t)$$

**Important Remark.** Converse of (*) is not true, *i.e.*,

$\phi_{X_1+X_2}(t) = \phi_{X_1}(t)\ \phi_{X_2}(t)$ does not imply that $X_1$ and $X_2$ are independent.

For example, let $X_1$ be a standard Cauchy variate with p.d.f.

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty$$

Then $\phi_{X_1}(t) = e^{-|t|}$ (c.f. Chapter 8)

Let $X_2 \equiv X_1$, *i.e.*, $P(X_1 = X_2) = 1$. ...(**)

Then $\phi_{X_2}(t) = e^{-|t|}$

Now $\phi_{X_1+X_2}(t) = \phi_{2X_1}(t) = \phi_{X_1}(2t) = e^{-2|t|}$

$= \phi_{X_1}(t)\ \phi_{X_2}(t)$

*i.e.* (*) is satisfied but obviously $X_1$ and $X_2$ are not independent, because of (**).

In fact, (*) will hold even if we take $X_1 = aX$ and $X_2 = bX$, a and b being real numbers so that $X_1$ and $X_2$ are connected by the relation :

$$\frac{X_1}{a} = X = \frac{X_2}{b} \quad \Rightarrow \quad aX_2 = bX_1 .$$

As another example let us consider the joint p.d.f. of two random variables $X$ and $Y$ given by

$$f(x, y) = \frac{1}{4a^2}[1 + xy(x^2 - y^2)] ; \ |x| \le a, \ |y| \le a, \ a > 0$$

$$= 0, \ \text{elsewhere}$$

Then the marginal p.d.f.'s of $X$ and $Y$ are given by

$$g(x)=\int_{-a}^{a} f(x,y)\,dy=\frac{1}{2a}\,;\ |x|\le a \qquad \text{(on simplification)}$$

$$h(y)=\int_{-a}^{a} f(x,y)\,dx=\frac{1}{2a}\,;\ |y|\le a$$

Then

$$\phi_X(t)=\int_{-\infty}^{\infty} e^{itx} g(x)\,dx=\frac{1}{2a}\int_{-a}^{a} e^{itx}\,dx$$

$$=\frac{e^{iat}-e^{-iat}}{2ait}=\frac{\sin at}{at}$$

Similarly $\qquad \phi_Y(t)=\dfrac{\sin at}{at}$

$$\therefore \qquad \phi_X(t)\ \phi_Y(t)=\left(\frac{\sin at}{at}\right)^2 \qquad \ldots(*)$$

The p.d.f. $k(z)$ of the random variable $Z=X+Y$ is given by the convolution of p.d.f.'s of $X$ and $Y$, *viz.*,

$$k(z)=\int f(u,z-u)\,du$$

$$=\frac{1}{4a^2}\int\left[1+u(z-u)\{u^2-(z-u)^2\}\right]du$$

$$=\frac{1}{4a^2}\int(1+3z^2u^2-2zu^3-z^3u)\,du,$$

the limits of integration for $u$ being in terms of $z$ and are given by (left as an exercise to the reader)

$$-a\le u\le z+a\,;\ u\le 0$$

and

$$z-a\le u\le a\,;\ u>0$$

Thus

$$k(z)=\begin{cases}\dfrac{1}{4a^2}\displaystyle\int_{-a}^{z+a}(1+3z^2u^2-2zu^3-z^3u)\,du=\dfrac{2a+z}{4a^2}\,; & -2a\le z\le 0\\[2ex] \dfrac{1}{4a^2}\displaystyle\int_{z-a}^{a}(1+3z^2u^2-2zu^3-z^3u)\,du=\dfrac{2a-z}{4a^2}\,; & 0<z\le 2a\\[2ex] 0,\ \text{elsewhere}\end{cases}$$

Now

$$\phi_{X+Y}(t)=\phi_Z(t)=\int_{-2a}^{2a} e^{itz}\,k(z)\,dz$$

$$=\int_{-2a}^{0}\left(\frac{2a+z}{4a^2}\right)e^{itz}\,dz+\int_{0}^{2a}\left(\frac{2a-z}{4a^2}\right)e^{itz}\,dz$$

$$= \int_0^{2a} \left(e^{-itz} + e^{itz}\right)\left(\frac{2a - z}{4a^2}\right) . dz$$

[Changing $z$ to $-z$ in the first integral]

$$= \frac{1}{2a^2} \int_0^{2a} (2a - z) \cos tz \, dz$$

$$= \frac{2 - 2\cos 2at}{4a^2t^2} = \frac{1 - \cos 2at}{2a^2t^2} \qquad \text{(on simplification)}$$

$$= \left(\frac{\sin at}{at}\right)$$

$$= \phi_X(t) . \phi_Y(t) \qquad \text{[From (*)]}$$

But $\quad g(x) . h(y) \neq f(x, y)$

$\Rightarrow \quad X$ and $Y$ are not independent.

However

$$\phi_{X_1, X_2}(t_1, t_2) = E(e^{it_1X_1 + it_2X_2}) = \phi_{X_1}(t) . \phi_{X_2}(t)$$

implies that $X_1$ and $X_2$ are independent.

*(For proof see Theorem 6·28)*

**Theorem 6·24.** *Effect of Change of Origin and Scale on Characteristic Function. If $U = \dfrac{X - a}{h}$, a and h being constants, then*

$$\phi_U(t) = e^{-iat/h} \phi_X(t/h)$$

In particular if we take $a = E(X) = \mu$ (say) and $h = \sigma_X = \sigma$ then the characteristic function of the standard variate

$$Z = \frac{X - E(X)}{\sigma_X} = \frac{X - \mu}{\sigma},$$

is given by $\quad \phi_Z(t) = e^{-i\mu t/\sigma} \phi_X(t/\sigma) \qquad$ ...(6·66)

**Definition.** *A random variable $X$ is said to be a Lattice variable or be lattice distributed, if for some $h > 0$,*

$$P\left[\frac{X}{h} \text{ is an integer}\right] = 1,$$

*$h$ is called a mesh.*

**Theorem 6·25.** *If $|\phi_X(s)| = 1$ for some $s \neq 0$, then for some real $a$, $X - a$ is a Lattice variable with mesh $h = 2\pi/|s|$.*

**Proof.** Consider any fixed $t$. We can write

$\phi_X(t) = |\phi_X(t)| \, e^{iat}$, *(a dependent on t), since any complex number z can be written as* $z = |z| e^{i\theta}$.

$$\therefore \quad |\phi_X(t)| = e^{-iat} \phi_X(t) = \phi_{(X-a)}(t)$$

$$= E[\cos t(X - a) + i \sin t(X - a)] = E[\cos t(X - a)]$$

since left-hand side being real, we must have $E[\sin t(X - a)] = 0$.

$$\therefore \quad 1 - |\phi_X(t)| = E[1 - \cos t(X - a)] \qquad ...(*)$$

If $|\phi_X(s)| = 1, s \neq 0$ then for some $a$ dependent on $s$, we have from (*)

$$E[1 - \cos s(X - a)] = 0 \qquad \text{...(**)}$$

But since $1 - \cos s(X - a)$ is a non-negative random variable, (**)

$\Rightarrow \quad P[1 - \cos s(X - a) = 0] = 1$

$\Rightarrow \quad P[\cos s(X - a) = 1] = 1$

$\Rightarrow \quad P[s(X - a) = 2n\pi] = 1$

$\Rightarrow \quad P\left[(X - a) = \dfrac{2n\pi}{|s|}\right] = 1$, for some $n = 0, 1, 2, \ldots$

Thus $(X - a)$ is a Lattice variable with mesh $h = \dfrac{2\pi}{|s|}$.

**6·12·3. Necessary and Sufficient Conditions for a Function $\phi(t)$ to be Characteristic Function.** Properties *(i)* to *(iv)* in § 6·12·1 are merely the necessary conditions for a function $\phi(t)$ to be the characteristic function of an r.v. $X$. Thus if a function $\phi(t)$ does not satisfy any one of these four conditions, it cannot be the characteristic function of an r.v. $X$. For example, the function

$$\phi(t) = \log(1 + t),$$

cannot be the c.f. of r.v. $X$ since $\phi(0) = \log 1 = 0 \neq 1$.

These conditions are, however, not sufficient. It has been shown (*c.f.* Methods of Mathematical Statistics by H. Cramer) that if $\phi(t)$ is near $t = 0$ of the form,

$$\phi(t) = 1 + 0(t^{2} + \delta), \delta > 0 \qquad \text{...(*)}$$

where $0(t^r)$ divided by $t^r$ tends to zero as $t \to 0$, then $\phi(t)$ cannot be the characteristic function unless it is identically equal to one. Thus, the functions

*(i)* $\phi(t) = e^{-t^4} = 1 + 0(t^4)$

*(ii)* $\phi(t) = \dfrac{1}{1 + t^4} = 1 + 0(t^4)$

being of the form (*) are not characteristic functions, though both satisfy all the necessary conditions.

We give below a set of sufficient but not necessary conditions, due to Polya for a function $\phi(t)$ to be the characteristic function :

$\phi(t)$ is a characteristic function if

(1) $\phi(0) = 1$,

(2) $\phi(t) = \phi(-t)$

(3) $\phi(t)$ is continuous

(4) $\phi(t)$ is convex for $t > 0$, *i.e.*, for $t_1, t_2 > 0$,

$$2\phi\left[\tfrac{1}{2}(t_1 + t_2)\right] \leq \phi(t_1) + \phi(t_2)$$

(5) $\lim\limits_{t \to \infty} \phi(t) = 0$.

Hence by Polya's conditions the functions $e^{-|t|}$ and $[1 + |t|]^{-1}$ are characteristic functions. However, Polya's conditions are only sufficient and not necessary for a characteristic function. For example, if $X \sim N(\mu, \sigma^2)$,

$$\phi(t) = e^{it\mu - t^2\sigma^2/2} \qquad [c.f. \text{ § } 8{\cdot}5]$$

and $\quad \phi(-t) \neq \phi(t)$.

Various necessary and sufficient conditions are known, the simplest seems to be the following, due to Cramer.

"In order that a given, bounded and continuous function $\phi(t)$ should be the characteristic function of a distribution, it is necessary and sufficient that $\phi(0) = 1$ and that the function

$$\phi(x, A) = \int_0^A \int_0^A \phi(t-u)\, e^{(t-u)ix} dt\, du$$

is real and non-negative for all real $x$ and all $A > 0$.

**6·12·4. Multi-variate Characteristic Function.** Then

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \text{ and } t = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}, \; t \text{ real}$$

be $z \times 1$ column vectors. Then characteristic function of $X$ is defined as

$$\phi_X(t) = E(e^{it'X}) = E[e^{i(t_1X_1 + t_2X_2 + \dots + t_nX_n)}] \qquad \dots(6{\cdot}67)$$

We may also write it as

$$\phi_{X_1, X_2, \dots X_n}(t_1, t_2, \dots, t_n) \text{ or } \phi_X(t_1, t_2, \dots, t_n)$$

**Some Properties.**

*(i)* $\phi_X(0, 0, \dots 0) = 1$

*(ii)* $\phi_{-X}(t) = \overline{\phi_X(t)}$

*(iii)* $|\phi_X(t)| \leq 1$

*(iv)* $\phi_X(t)$ is uniformly continuous in $n$-dimensional Euclidian space.

*(v)* If $f_X(\cdot)$ is p.d.f. of X,

$$\phi_X(t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_X(x_1, x_2, \dots, x_n)\, e^{i \Sigma t_j x_j}\, dx_1\, dx_2 \dots dx_n$$

*(vi)* $\phi_X(t) = \phi_Y(t)$ for all t, then $X$ and $Y$ have the same distribuiton.

*(vii)* If $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \phi_X(t_1, t_2, \dots t_n) \right| dt_1\, dt_2 \dots dt_n < \infty$,

then $X$ is absolutely continuous and has a uniformly continuous p.d.f.

$$f_X(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\Sigma t_j x_j}\, \phi_X(t_1, t_2, \dots, t_n)\, dt_1\, dt_2 \dots dt_n$$

*(viii)* The random variables $X_1, X_2, \dots, X_n$ are (mutually) independent iff

$$\phi_{X_1, X_2, \dots X_n}(t_1, t_2, \dots, t_n) = \phi_{X_1}(t_1)\, \phi_{X_2}(t_2) \dots \phi_{X_n}(t_n)$$

**Remark.** *Multivariate Moment Generating Function.* Similarly, the m.g.f. of vector $X = (X_1, X_2, \dots, X_n)'$ is given by :

$$M_X(t) = E(e^{t'X}) = E(e^{t_1X_1 + t_2X_2 + \ldots + t_nX_n}) \qquad \ldots(6\cdot68)$$

We may also write :

$$M_X(t) = M_{X_1, X_2, \ldots X_n}(t_1, t_2, \ldots, t_n) = E(e^{t_1X_1 + t_2X_2 + \ldots + t_nX_n})$$

In particular, for two variates $X_1$ and $X_2$

$$M_X(t) = M_{X_1, X_2}(t_1, t_2) = E(e^{t_1X_1 + t_2X_2}) = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{t_1^r}{r!}\frac{t_2^s}{s!} E(X_1^r X_2^s), \qquad \ldots(6\cdot69)$$

provided it exists for $-h_1 < t_1 < h_1$ and $-h_2 < t_2 < h_2$, where $h_1$ and $h_2$ are positive.

$$M_{X_1, X_2}(t_1, 0) = E(e^{t_1X_1}) = M_{X_1}(t_1) \qquad \ldots(6\cdot69a)$$

$$M_{X_1, X_2}(0, t_2) = E(e^{t_2X_2}) = M_{X_2}(t_2) \qquad \ldots(6\cdot69b)$$

If $M(t_1, t_2)$ exists, the moments of all orders of $X$ and $Y$ exist and are given by:

$$E(X_2^r) = \left[\frac{\partial^r M(t_1, t_2)}{\partial t_2^r}\right]_{t_1 = t_2 = 0} = \frac{\partial^r M(0, 0)}{\partial t_2^r} \qquad \ldots(6\cdot70)$$

$$E(X_1^r) = \left[\frac{\partial^r M(t_1, t_2)}{\partial t_1^r}\right]_{t_1 = t_2 = 0} = \frac{\partial^r M(0, 0)}{\partial t_1^r} \qquad \ldots(6\cdot70a)$$

$$E(X_1^r X_2^s) = \left[\frac{\partial^{r+s} M(t_1, t_2)}{\partial t_1^r \partial t_2^s}\right]_{t_1 = t_2 = 0} = \frac{\partial^{r+s} M(0, 0)}{\partial t_1^r \partial t_2^s} \qquad \ldots(6\cdot70b)$$

Cumulant generating function of $X = (X_1, X_2)'$ is given by :

$$K_{X_1, X_2}(t_1, t_2) = \log M_{X_1, X_2}(t_1, t_2). \qquad \ldots(6\cdot71)$$

**Example 6·42.** *For a distribution, the cumulants are given by*

$$\kappa_r = n\,[(r-1)\,!], \quad n > 0$$

*Find the characteristic function.* **(Delhi Univ. B.Sc. (Stat. Hons.), 1990)**

**Solution.** The cumulant generating function $K(t)$, if it exists, is given by

$$K(t) = \sum_{r=1}^{\infty} \frac{(it)^r}{r!}\kappa_r = \sum_{r=1}^{\infty} \frac{(it)^r}{r!} n\{(r-1)\,!\} = n\sum_{r=1}^{\infty} \frac{(it)^r}{r}$$

$$= n\left[it + \frac{(it)^2}{2} + \frac{(it)^3}{3} + \ldots\right] = n\,[\log(1 - it)]$$

$$= -n\log(1 - it) = \log(1 - it)^{-n}$$

Also we have

$$K(t) = \log\phi(t) = \log(1 - it)^{-n}$$

$$\therefore \qquad \phi(t) = (1 - it)^{-n}$$

**Remark.** This is the characteristic function of the gamma distribution : (*c.f.* § 8·3·1)

$$f(x) = \frac{e^{-x}\, x^{n-1}}{\Gamma(n)}\ ;\ n > 0,\ 0 < x < \infty.$$

**Example 6·43.** *The moments about origin of a distribution are given by*

$$\mu_r' = \frac{\Gamma(\nu + r)}{\Gamma(\nu)}.$$

*Find the characteristic function.*

**(Madurai Kamaraj Univ. B.Sc., 1990)**

**Solution.** We have

$$\phi(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{r\,!}\mu_r' = \sum_{r=0}^{\infty} \frac{(it)^r}{r\,!} \cdot \frac{\Gamma(\nu + r)}{\Gamma(\nu)}$$

$$= \sum_{r=0}^{\infty} \frac{(it)^r}{r\,!} \cdot \frac{(\nu + r - 1)\,!}{(\nu - 1)\,!}$$

$$= \sum_{r=0}^{\infty} (it)^r \cdot {}^{\nu + r - 1}C_r = \sum_{r=0}^{\infty} (-1)^r \cdot {}^{-\nu}C_r\, (it)^r$$

$$[\because\ {}^{-\nu}C_r = (-1)^r \cdot {}^{\nu + r - 1}C_r \ \Rightarrow\ (-1)^r \cdot {}^{-\nu}C_r = {}^{\nu + r - 1}C_r]$$

$$\therefore \qquad \phi(t) = \sum_{r=0}^{\infty} {}^{-\nu}C_r\, (-it)^r = (1 - it)^{-\nu}$$

**Example 6·44.** *Show that*

$$e^{itx} = 1 + (e^{it} - 1)\, x^{(1)} + (e^{it} - 1)^2 \frac{x^{(2)}}{2\,!} + \ldots + (e^{it} - 1)^r \cdot \frac{x^{(r)}}{r\,!} + \ldots$$

*where* $x^{(r)} = x(x-1)(x-2)\ldots(x-r+1)$. *Hence show that*

$\mu_{(r)}' = [D^r\, \phi(t)]_{t=0}$, *where* $D = \dfrac{d}{d(e^{it})}$ *and* $\mu_{(r)}'$ *is the* $r^{th}$ *factorial moment.*

**Solution.** We have

$$\text{R.H.S.} = 1 + (e^{it} - 1)\, x^{(1)} + (e^{it} - 1)^2 \cdot \frac{x^{(2)}}{2\,!} + \ldots + (e^{it} - 1)^r \cdot \frac{x^{(r)}}{r\,!} + \ldots$$

$$= 1 + (e^{it} - 1)({}^xC_1) + (e^{it} - 1)^2({}^xC_2)$$
$$+ (e^{it} - 1)^3({}^xC_3) + \ldots + (e^{it} - 1)^r({}^xC_r)$$

$$= [1 + (e^{it} - 1)]^x = e^{itx} = \text{L.H.S.}$$

By def.

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx$$

$$= \int_{-\infty}^{\infty} \left[ 1 + (e^{it} - 1)x^{(1)} + (e^{it} - 1)^2 \cdot \frac{x^{(2)}}{2!} + \cdots + (e^{it} - 1)^r \cdot \frac{x^{(r)}}{r!} + \ldots \right] f(x)\, dx$$

$$= 1 + (e^{it} - 1) \int_{-\infty}^{\infty} x^{(1)} f(x)\, dx + \frac{(e^{it} - 1)^2}{2!} \int_{-\infty}^{\infty} x^{(2)} f(x)\, dx + \ldots$$

$$+ \frac{(e^{it} - 1)^r}{r!} \int_{-\infty}^{\infty} x^{(r)} f(x)\, dx + \ldots$$

$$\therefore \qquad [D^r \phi(t)]_{t=0} = \left[ \frac{d^r \phi(t)}{d(e^{it})^r} \right]_{t=0} = \int_{-\infty}^{\infty} x^{(r)} f(x)\, dx = \mu_{(r)}'$$

where $\mu_{(r)}'$ is the $r$th factorial moment.

**Theorem 6·26. (Inversion Theorem).** **Lemma.** *If* $(a-h, a+h)$ *is the continuity interval of the distribution function* $F(x)$, *then*

$$F(a+h) - F(a-h) = \lim_{T \to \infty} \frac{1}{\pi} \int_{-T}^{T} \frac{\sin ht}{t} e^{-ita} \phi(t)\, dt,$$

$\phi(t)$ *being the characteristic function of the distribution.*

**Corollary.** If $\phi(t)$ is absolutely integrable over $R^1$, *i.e.*, if

$$\int_{-\infty}^{\infty} |\phi(t)|\, dt < \infty,$$

then the derivative of $F(x)$ exists, which is bounded, continuous on $R^1$ and is given by

$$f(x) = F'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi(t)\, dt, \qquad \ldots(6·72)$$

for every $x \in R^1$.

**Proof.** In the above lemma replacing $a$ by $x$ and on dividing by $2h$, we have

$$\frac{F(x+h) - F(x-h)}{2h} = \frac{1}{2\pi} \cdot \lim_{T \to \infty} \int_{-T}^{T} \frac{\sin ht}{ht} e^{itx} \phi(t)\, dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\sin ht}{ht} e^{-itx} \phi(t)\, dt$$

$$\therefore \lim_{h \to 0} \frac{F(x+h) - F(x-h)}{2h} = \frac{1}{2\pi} \lim_{h \to 0} \int_{-\infty}^{\infty} \frac{\sin ht}{ht} \cdot e^{-itx} \phi(t)\, dt$$

Since $$\int_{-\infty}^{\infty} |\phi(t)|\, dt < \infty,$$

the integrand on the right hand side is bounded by an integrable function and hence by Dominated Convergence Theorem, we get

$$\lim_{h \to 0} \frac{F(x+h) - F(x-h)}{2h} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{h \to 0} \left( \frac{\sin ht}{ht} \right) \cdot e^{-itx} \phi(t)\, dt$$

By mean value theorem of differential calculus, we have

$$\lim_{h \to 0} \frac{F(x+h) - F(x-h)}{2h} = F'(x) = f(x),$$

where $f(.)$ is the p.d.f. corresponding to $\phi(t)$. Thus

$$f(x) = F'(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\phi(t)\,dt,$$

as desired.

**Remark.** Consider the function $\mathcal{F}_c$ defined by

$$\mathcal{F}_c = \int_{-c}^{c} e^{-itx}\,\phi(t)\,dt$$

Now if $F'(x) = f(x)$ exists, then

$$\lim_{c \to \infty} \frac{\mathcal{F}_c}{2c} = \lim_{c \to \infty} \frac{1}{2c}\int_{-c}^{c} \phi(t)\,e^{-itx}\,dt$$

$$= \lim_{c \to \infty} \left\{\frac{1}{2c}\,.\,2\pi f(x)\right\} = 0$$

Hence $\dfrac{\mathcal{F}_c}{2c} \to 0$ at all points where $F(x)$ is continuous. In other words, if the prbability distribution is continuous

$$\frac{\mathcal{F}_c}{2c} \to 0 \text{ as } c \to \infty$$

If, however, the frequency function is discontinuous, *i.e.*, distribution is discrete, consider one point of discontinuity say, the frequency $f_j$ at $x = x_j$. Then the contribution of $x_j$ to $\phi(t)$ is $f_j e^{itx_j}$ and hence its contributions to $\mathcal{F}_c$ will be

$$\int_{-c}^{c} f_j e^{itx_j}\,e^{-itx}\,dt$$

$$\therefore \quad \lim_{c \to \infty} \frac{\mathcal{F}_c}{2c} = \lim_{c \to \infty} \frac{1}{2c} f_j \int_{-c}^{c} e^{it(x_j - x)}\,dt$$

$$= \lim_{c \to \infty} \frac{1}{2c} f_j \left[\frac{e^{it(x_j - x)}}{i(x_j - x)}\right]_{-c}^{c}$$

$$= \begin{cases} 0 & \text{for } x \neq x_j \\ f_j & \text{for } x = x_j \end{cases}$$

*Hence if $\mathcal{F}_c/2c \to 0$ at a point, there is no discontinuity in the distribution function at that point, but if it tends to a positive number $f_j$, the distribution is discontinuous at that point and the frequency is $f_j$. This gives us a criterion whether a given characteristic function represents a continuous distribution or not.*

**Theorem 6·27. Uniqueness Theorem of Characteristic Functions.** *Characteristic function uniquely determines the distribution, i.e., a necessary and sufficient condition for two distributions with p.d.f.'s $f_1(.)$ and $f_2(.)$ to be identical is that their characteristic functions $\phi_1(t)$ and $\phi_2(t)$ are identical.*

**Proof.** If $f_1(\cdot) = f_2(\cdot)$, then from the definition of characteristic function, we get

$$\phi_1(t) = \int_{-\infty}^{\infty} e^{itx} f_1(x)\, dx = \int_{-\infty}^{\infty} e^{itx} f_2(x)\, dx = \phi_2(t)$$

Conversely if $\phi_1(t) = \phi_2(t)$, then from corollary to Theorem 6·26, we get

$$f_1(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_1(t)\, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_2(t)\, dt = f_2(x)$$

**Remark.** This is one of the most fundamental theorems in the distribution theory. It implies that corresponding to a distribution there is only one characteristic function and corresponding to a given characteristic function, there is only one distribution. This *one to one correspondence between characteristic functions and the p.d.f.'s enables us to identify the form of the p.d.f. from that of characteristic function.*

**Theorem 6·28.** *Necessary and sufficient condition for the random variables $X_1$ and $X_2$ to be independent is that their joint characteristic function is equal to the product of their individual characteristic functions, i.e.,*

$$\phi_{X_1, X_2}(t_1, t_2) = \phi_{X_1}(t_1)\, \phi_{X_2}(t_2) \qquad \ldots(*)$$

**Proof.** (*i*) *Condition is Necessary.* If $X_1$ and $X_2$ are independent then we have to show that (*) holds. By def.,

$$\begin{aligned} \phi_{X_1 X_2}(t_2, t_2) &= E\left(e^{it_1X_1 + it_2X_2}\right) = E\left(e^{it_1X_1} \cdot e^{it_2X_2}\right) \\ &= E\left(e^{it_1X_1}\right) E\left(e^{it_2X_2}\right) (\because X_1, X_2 \text{ are independent}) \\ &= \phi_{X_1}(t_1)\, \phi_{X_2}(t_2), \end{aligned}$$

as required.

(*ii*) *Condition is sufficient.* We have to show that if (*) holds, then $X_1$ and $X_2$ are independent.

Let $f_{X_1, X_2}(x_1, x_2)$ be the joint p.d.f. of $X_1$ and $X_2$ and $f_1(x_1)$ and $f_2(x_2)$ be the marginal p.d.f.'s of $X_1$ and $X_2$ respectively. Then by definition (for continuous r.v.'s), we get

$$\phi_{X_1}(t_1) = \int_{-\infty}^{\infty} e^{it_1 x_1} f_1(x_1)\, dx_1$$

$$\phi_{X_2}(t_2) = \int_{-\infty}^{\infty} e^{it_2 x_2} f_2(x_2)\, dx_2$$

$$\begin{aligned} \therefore \quad \phi_{X_1}(t_1)\, \phi_{X_2}(t_2) &= \left[ \int_{-\infty}^{\infty} e^{it_1 x_1} f_1(x_1)\, dx_1 \right] \left[ \int_{-\infty}^{\infty} e^{it_2 x_2} f_2(x_2)\, dx_2 \right] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(t_1 x_1 + t_2 x_2)} f_1(x_1) f_2(x_2)\, dx_1\, dx_2 \qquad \ldots(**) \end{aligned}$$

by Fubini's theorem, since the integrand is bounded by an integrable function.

Also by def.

$$\phi_{X_1, X_2}(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(t_1x_1 + t_2x_2)} f(x_1, x_2)\, dx_1\, dx_2$$

If (*) holds, we get from (**)

$$\phi_{X_1, X_2}(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(t_1x_1 + t_2x_2)} f_1(x_1) f_2(x_2)\, dx_1\, dx_2$$

Hence by uniqueness theorem of characteristic functions, we get

$f(x_1, x_2) = f_1(x_1) f_2(x_2)$,

which implies that $X_1$ and $X_2$ are independent.

**Remarks. 1.** For discrete r.v.'s, the result is established by using summation instead of integration.

**2.** The result can be generalised to the case of more than two variables.

Necessary and sufficient condition for the mutual independence of random variables $X_i$, $(i = 1, 2, 3, \ldots, n)$ is that

$$\phi_{X_1, X_2, \ldots, X_n}(t_1, t_2, \ldots, t_n) = \phi_{X_1}(t_1)\, \phi_{X_2}(t_2) \ldots \phi_{X_n}(t_n)$$

**3.** In terms of moment generating functions, the necessary and sufficient condition for the r.v.'s $X_1, X_2, \ldots, X_n$ to be mutually independent is that

$$M_{X_1, X_2, \ldots, X_n}(t_1, t_2, \ldots, t_n) = M_{X_1}(t_1)\, M_{X_2}(t_2) \ldots M_{X_n}(t_n)$$

provided m.g.f.'s exist.

**Theorem 6·29. Hally-Bray Theorem.** *If the sequence of distribution functions $\{F_n(x)\}$ converges to the distribution function $F(x)$ at all the points of continuity of the latter and $g(x)$ is bounded continuous function over the line $R^1(-\infty, \infty)$, then*

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} g(x)\, dF_n(x) = \int_{-\infty}^{\infty} g(x)\, dF(x) \qquad \ldots(*)$$

**Corollary.** If $F_n(x) \to F(x)$, then the corresponding sequence of characteristic functions $\phi_n(t)$ of $F_n(x)$ converges to the characteristic function $\phi(t)$ of $F$ at every point '$t$'.

**Proof.** $\cos tx$ and $\sin tx$ are continuous and bounded functions of $x$ for all $t$ and hence from (*), we get

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} \cos tx\; dF_n(x) = \int_{-\infty}^{\infty} \cos tx\; dF(x)$$

and

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} \sin tx\; dF_n(x) = \int_{-\infty}^{\infty} \sin tx\; dF(x)$$

$$\therefore \lim_{n \to \infty} \int_{-\infty}^{\infty} (\cos tx + i \sin tx)\; dF_n(x) = \int_{-\infty}^{\infty} (\cos tx + i \sin tx).$$

$$\Rightarrow \quad \lim_{n \to \infty} \int_{-\infty}^{\infty} e^{itx}\, dF_n(x) = \int_{-\infty}^{\infty} e^{itx}\, dF(x).$$

$$\Rightarrow \quad \phi_n(t) \to \phi(t) \text{ as } n \to \infty$$

**Theorem 6·30. Continuity Theorem for Characteristic Functions.** For a sequence of distribution functions $\{F_n(x)\}$ with the corresponding sequence of characteristic functions $\{\phi_n(t)\}$, a necessary and sufficient condition that $F_n(x) \to F(x)$ at all points of continuity of $F$ is that for every real $t$, $\phi_n(t) \to \phi(t)$, which is continuous at $t = 0$ and $\phi(t)$ is the characteristic function corresponding to $F$.

**Example 6·45.** *Let $F_n(x)$ be the distribution function defined by*

$$\begin{aligned} F_n(x) &= 0 \text{ for } x \leq -n \\ &= \frac{x+n}{2n} \text{ for } -n < x < n \\ &= 1 \text{ for } x \geq n \end{aligned}$$

*Is the limit $F_n(x)$ a distribution function? If not, why?*

**Solution.**
$$\phi_n(t) = \int_{-n}^{n} e^{itx} \frac{1}{2n}\, dx \qquad (\because f_n(x) = F_n'(x))$$

$$= \frac{1}{2n}\left[\frac{e^{itn} - e^{-itn}}{it}\right] = \frac{\sin nt}{nt}$$

$$\phi(t) = \lim_{n \to \infty} \phi_n(t) = \lim_{n \to \infty} \frac{\sin nt}{nt}$$

$$= \begin{cases} 1 & \text{if } t = 0 \\ 0 & \text{if } t \neq 0 \end{cases}$$

*i.e.*, $\phi(t)$ is discontinuous at $t = 0$.

and
$$\lim_{n \to \infty} F_n(x) = \frac{1}{2}$$

Hence $F(x)$ is not a distribution function.

**Example 6·46.** *Find the distribution for which characteristic function is* (*a*) $\phi(t) = (q + pe^{it})^n$, (*b*) $\phi(t) = e^{-t^2\sigma^2/2}$

**Solution.** (*a*)
$$\phi(t) = (q + pe^{it})^n = \sum_{j=0}^{n} {}^nC_j\, p^j q^{n-j} e^{itj}$$

We have

$$\mathcal{F}_c = \int_{-c}^{c} e^{-itx}\, \phi(t)\, dt = \int_{-c}^{c} \left\{ e^{-itx} \sum_{j=0}^{n} {}^nC_j\, p^j q^{n-j} e^{itj} \right\} dt$$

$$= \sum_{j=0}^{n} \left[ {}^nC_j p^j q^{n-j} \int_{-c}^{c} e^{-it(x-j)} dt \right]$$

(*i*) If $x \neq j$,

$$\mathcal{F}_c = \sum_{j=0}^{n} {}^nC_j p^j q^{n-j} \left| \frac{e^{-it(x-j)}}{-i(x-j)} \right|_{-c}^{c} = \sum_{j=0}^{n} {}^nC_j p^j q^{n-j} \left[ \frac{e^{ic(x-j)} - e^{-ic(x-j)}}{i(x-j)} \right]$$

$$= \sum_{j=0}^{n} \left[ {}^nC_j p^j q^{n-j} . \frac{2i \sin \{c(x-j)\}}{(x-j)} \right]$$

$$\therefore \qquad \lim_{c \to \infty} \frac{\mathcal{F}_c}{2c} \to 0 \ \forall x.$$

Hence there is no discontinuity in the distribution function when $x \neq j$.

(*ii*) If $x = j$,

$$\mathcal{F}_c = \sum_{j=0}^{n} \left[ {}^nC_j p^j q^{n-j} \int_{-c}^{c} dt \right] = 2c \sum_{j=0}^{n} {}^nC_j p^j q^{n-j} = 2c\,(q+p)^n = 2c$$

Since $\dfrac{\mathcal{F}_c}{2c} \to 1$ at $x = j$, the distribution function is discontinuous and its frequency is ${}^nC_j p^j q^{n-j}$.

(*b*) Let $\displaystyle \mathcal{F}_c = \int_{-c}^{c} e^{-itx - \frac{1}{2}\sigma^2 t^2} dt$

$$\therefore \qquad |\mathcal{F}_c| \leq \int_{-c}^{c} \left| e^{-itx - \frac{1}{2}\sigma^2 t^2} \right| dt \leq \int_{-c}^{c} e^{-\frac{1}{2}\sigma^2 t^2} dt$$

$$\leq \int_{-\infty}^{\infty} e^{-\frac{1}{2}\sigma^2 t^2} dt = \frac{\sqrt{2\pi}}{\sigma}$$

$$\therefore \qquad \lim_{c \to \infty} \frac{\mathcal{F}_c}{2c} = 0 \ \forall x.$$

Hence the distribution function is continuous for all $x$.

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} e^{-t^2 \sigma^2/2} dt$$

$$= \frac{1}{2\pi} e^{-x^2/2\sigma^2} \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} \left( t\sigma + \frac{ix}{\sigma} \right)^2 \right\} dt$$

Put $\quad t\sigma + \dfrac{ix}{\sigma} = \xi$, *i.e.*, $\sigma dt = d\xi$

$$\therefore f(x)=\frac{1}{2\pi}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}\int_{-\infty}^{\infty} e^{-\xi^2/2}\,\frac{d\xi}{\sigma}$$

Hence

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

which is the p.d.f. of normal distribution.

**Example 6·47.** *Find the density function $f(x)$ corresponding to the characteristic function defined as follows :*

$$\phi(t)=\begin{cases}1-|t|, & |t|\le 1\\ 0, & |t|>1\end{cases}$$

**[Delhi Univ, B.Sc. (Maths Hons.), 1989]**

**Solution.** By Inversion Theorem, the p.d.f. of $X$ is given by :

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\phi(t)\,dt$$

$$=\frac{1}{2\pi}\int_{-1}^{0} e^{-itx}(1+t)\,dt+\frac{1}{2\pi}\int_{0}^{1} e^{-itx}(1-t)\,dt\ *$$

Now

$$\int_{-1}^{0} e^{-itx}(1+t)\,dt=\left[\frac{e^{-itx}}{-ix}(1+t)\right]_{-1}^{0}+\frac{1}{ix}\int_{-1}^{0} e^{-itx}\,dt$$

$$=-\frac{1}{ix}+\frac{1}{ix}\left[\frac{e^{-itx}}{-ix}\right]_{-1}^{0}$$

$$=-\frac{1}{ix}+\frac{1}{(ix)^2}(e^{ix}-1)$$

Similarly,

$$\int_{0}^{1} e^{-itx}(1-t)\,dt=\frac{1}{ix}+\frac{1}{(ix)^2}(e^{-ix}-1)$$

$$\therefore\qquad f(x)=\frac{1}{2\pi}\left[\frac{1}{(ix)^2}\left\{e^{ix}-1+e^{-ix}-1\right\}\right]$$

$$=\frac{1}{\pi x^2}\left[1-\frac{e^{ix}+e^{-ix}}{2}\right]=\frac{1}{\pi}\cdot\frac{1-\cos x}{x^2},\ -\infty<x<\infty$$

**EXERCISE 6 (c)**

1. Define m.g.f. of a random variable. Hence or otherwise find the m.g.f. of:

---

* $\because$ for $-1<t<0$, $|t|=-t$ and for $0<t<1$, $|t|=+t$

$$(i)\ Y = aX + b, \quad (ii)\ Y = \frac{X - m}{\sigma}.$$

**[Sri Venkat Univ. B.Sc., Sept. 1990; Kerala Univ. B.Sc., Sept. 1992]**

**2.** The random variable $X$ takes the value $n$ with probability $1/2^n$, $n = 1, 2, 3, \ldots$ Find the moment generating function of $X$ and hence find the mean and variance of $X$.

**3.** Show that if $\overline{X}$ is mean of $n$ independent random variables, then

$$M_{\overline{X}}(t) = \left[M_X\left(\frac{t}{n}\right)\right]^n$$

**4.** (*a*) Define moments and moment generating function (m.g.f.) of a random variable $X$. If $M(t)$ is the m.g.f. of a random variable $X$ about the origin, show that the moment $\mu_r'$ is given by

$$\mu_r' = \left[\frac{d^r M(t)}{dt^r}\right]_{t=0}$$ **[Baroda Univ. B.Sc., 1992]**

(*b*) If $\mu_r'$ is the $r$th order moment about the origin and $\kappa_j$ is the cumulant of $j$th order, prove that

$$\frac{\delta \mu_r'}{\delta \kappa_j} = \binom{r-1}{j-1} \mu'_{r-j}$$

(*c*) If $\mu_r'$ is the $r$th moment about the origin of a variable $X$ and if $\mu_r' = r!$, find the m.g.f. of $X$.

**5.** (*a*) A random variable '$X$' has probability function

$$p(x) = \frac{1}{2^x};\ x = 1, 2, 3, \ldots$$

Find the M.G.F., mean and variance.

(*b*) Show that the m.g.f. of r.v. $X$ having the p.d.f.

$$f(x) = \tfrac{1}{3},\ -1 < x < 2$$
$$= 0, \text{ elsewhere,}$$

is $$M(t) = \frac{e^{2t} - e^{-t}}{3t},\ t \neq 0$$
$$= 1,\ t = 0$$ **[Gujarat Univ. B.Sc., Oct. 1991]**

(*c*) A random variable '$X$' has the density function :

$$f(x) = \frac{1}{2\sqrt{x}},\ 0 < x < 1$$
$$= 0, \text{ elsewhere}$$

Obtain the moment generating function and hence the mean and variance.

**6.** $X$ is a random variable and $p(x) = ab^x$, where $a$ and $b$ are positive, $a + b = 1$ and $x$ taking the values 0, 1, 2, ... Find the moment generating function of $X$. Hence show that

$$m_2 = m_1(2m_1 + 1)$$

$m_1$ and $m_2$ being the first two moments.

7. Find the characteristic function of the following distributions and their variances :

(i) $dF(x) = ae^{-ax}\,dx,\ (a > 0, x > 0)$

(ii) $dF(x) = \frac{1}{2}e^{-|x|}\,dx,\ (-\infty < x < \infty)$

(iii) $P(X = j)$ lineu$=\binom{n}{j} p^j q^{n-j},\ (0 < p < 1, q = 1 - p, j = 0, 1, 2, ..., n).$

8. Obtain the m.g.f. of the random variable $X$ having p.d.f.,

$$f(x) = \begin{cases} x, & \text{for } 0 \le x < 1 \\ 2 - x, & \text{for } 1 \le x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

Determine $\mu_1'$, $\mu_2$, $\mu_3$ and $\mu_4$.

Ans. $\left(\frac{e^t - 1}{t}\right)^2$, $\mu_1' = 1$, $\mu_2' = 7$.

9. (a) Define cumulants and obtain the first four cumulants in terms of central moments.

(b) If $X$ is a variable with zero mean and cumulants $\kappa_r$, show that the first two cumulants $l_1$ and $l_2$ of $X^2$ are given by $l_1 = \kappa_2$ and $l_2 = 2\kappa_2^2 + \kappa_4$.

10. Show that the $r$th cumulant for the distribution

$$f(x) = ce^{-cx}, \text{ where } c \text{ is positive and } 0 \le x < \infty$$

is $$\frac{1}{c^r}.(r-1)!$$

11. If $X$ is a random variable with cumulants $\kappa_r$ ; $r = 1, 2, ....$ Find the cumulants of

(i) $cX$, (ii) $c + X$, where $c$ is a constant.

12. (a) Define the characteristic function of a random variable. Show that the characteristic function of the sum of two independent variables is equal to the product of their characteristic functions.

(b) If $X$ is a random variable having cumulants $\kappa_r$ ; $r = 1, 2, ...$ given by

$$\kappa_r = (r-1)!\,pa^{-r} ; p > 0, a > 0,$$

find the characteristic function of $X$.

(c) Prove that the characteristic function of a random variable $X$ is real if and only if $X$ has a symmetric distribution about 0.

13. Define $\phi(t)$, the characteristic function of a random variable. Find the characteristic function of a random variable $X$ defined as follows :

$$f(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \le x \le 1 \\ 0, & x > 1 \end{cases}$$

Ans. $e^{(it-1)}/it$

14. For the joint distribution of two-dimensional random variable $(X, Y)$ given by

$$f(x, y) = \frac{1}{4a^2}\left[1 + xy\,(x^2 - y^2)\right], \; |x| \leq a\,; \; |y| \leq a, a > 0$$
$$= 0, \text{ elsewhere },$$

show that the characteristic function of $X + Y$ is equal to the product of the characteristic functions of $X$ and $Y$. Comment on the result.

**Hint.** See remark to Theorem 6·22, page 6·81.

15. Let $K(t_1, t_2) = \log_e M(t_1, t_2)$ where $M(t_1, t_2)$ is the m.g.f. of $X$ and $Y$. Show that :

$$\frac{\partial K(0,0)}{\partial t_1} = E(X)\,; \quad \frac{\partial^2 K(0,0)}{\partial t_1^2} = \text{Var}\, X\,; \quad \frac{\partial^2 K(0,0)}{\partial t_1\, \partial t_2} = \text{Cov}\,(X, Y)$$

**[Delhi Univ. B.A.(Stat. Hons.), Spl. Course 1987]**

## OBJECTIVE TYPE QUESTIONS

I. Comment on the following, giving examples, if possible :

*(i)* M.g.f. of a r.v. always exist.

*(ii)* Characteristic function of a r.v. always exists.

*(iii)* M.g.f. is not affected by change of origin or/and scale.

*(iv)* $\phi_{X+Y}(t) = \phi_X(t) \cdot \phi_Y(t)$ implies $X$ and $Y$ *are independent.*

*(v)* $\phi_X(t) = \phi_Y(t)$ implies $X$ and $Y$ have the same distribution.

*(vi)* $\phi(0) = 1$ and $|\phi(t)| \leq 1$.

*(vii)* Variance of a r.v. is 5 and its mean does not exist.

*(ix)* It is possible to find a r.v. whose first $k$ moments exist but $(k+1)^{th}$ moment does not exist.

*(x)* If a r.v. $X$ has a symmetrical distribution about origin then

*(a)* $\phi_X(t)$ is even valued function of $t$.

*(b)* $\phi_X(t)$ is complex valued function of $t$.

II. *(a)* Can the following be the characteristic functions of any distribution ? Give reasons.

*(i)* $\log(1 + t)$, *(ii)* $\exp(-t^4)$, *(iii)* $1/(1 + t^4)$.

*(b)* Prove that $\phi(t) = \exp(-t^{\alpha})$, cannot be a characteristic function unless $\alpha = 2$.

III. State the relations, if any, between the following :

*(i)* $E(X^r)$ and $\phi_X(t)$.

*(ii)* $M_X(t)$ and $M_{X-a}(t)$, $a$ being constant.

*(iii)* $M_X(t)$ and $M_{(X-a)/h}(t)$, $a$ and $h$ being constants.

*(iv)* $\phi_X(t)$ and p.d.f. of $X$.

*(v)* $\mu_r$ and $\mu_r'$.

*(vi)* First four cumulants in terms of first four moments about mean.

IV. Let $X_1, X_2, ..., X_n$ be $n$ *i.i.d.* (independent and identically distributed) r.v.'s with m.g.f. $M(t)$. Then prove that

$$M_{\bar{X}}(t) = [M(t/n)]^n,$$

where
$$\bar{X} = \sum_{i=1}^{n} X_i/n$$

V. If $X_1, X_2, ..., X_n$ are independent r.v.'s then prove that

$$M_{\sum_{i=1}^{n} c_i X_i}(t) = \prod_{i=1}^{n} M_{X_i}(c_i t).$$

VI. Fill in the blanks :

*(i)* If $\int_{-\infty}^{\infty} |\phi_X(t)|\,dt < \infty$, then p.d.f. of $X$ is given by ...

*(ii)* $X_1$ and $X_2$ are independent if and only if ... .
(Give result in terms of characteristic functions.)

*(iii)* If $X_1$ and $X_2$ are independent then

$$\phi_{X_1 - X_2}(t) = \ldots$$

*(iv)* $\phi(t)$ is ... defined and is ... for all $t$ in $(-\infty, \infty)$.

VII. Examine critically the following statements :

*(a)* Two distributions having the same set of moments are identical.

*(b)* The characteristic function of a certain non-degenerate distribution is $e^{-t^3}$.

**6·13. Chebychev's Inequality.** The role of standard deviation as a parameter to characterise variance is precisely interpreted by means of the well known Chebychev's inequality. The theorem discovered in 1853 was later on discussed in 1856 by Bienayme.

**Theorem 6·31.** *If $X$ is a random variable with mean $\mu$ and variance $\sigma^2$, then for any positive number $k$, we have*

$$P\{|X - \mu| \geq k\sigma\} \leq 1/k^2 \qquad \ldots(6{\cdot}73)$$

or
$$P\{|X - \mu| < k\sigma\} \geq 1 - (1/k^2) \qquad \ldots(6{\cdot}73\,a)$$

**Proof. Case (i).** *$X$ is a continuous r.v. By def.,*

$$\sigma^2 = \sigma_X^2 = E[X - E(X)]^2 = E[X - \mu]^2$$

$$= \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\,dx, \text{ where } f(x) \text{ is p.d.f. of } X.$$

$$= \int_{-\infty}^{\mu - k\sigma} (x-\mu)^2 f(x) dx + \int_{\mu - k\sigma}^{\mu + k\sigma} (x-\mu)^2 f(x)\, dx + \int_{\mu + k\sigma}^{\infty} (x-\mu)^2 f(x)\, dx$$

$$\geq \int_{-\infty}^{\mu - k\sigma} (x-\mu)^2 f(x)\, dx + \int_{\mu + k\sigma}^{\infty} (x-\mu)^2 f(x)\, dx \qquad \ldots(*)$$

We know that :

$$x \leq \mu - k\sigma \text{ and } x \geq \mu + k\sigma \quad \Leftrightarrow \quad |x - \mu| \geq k\sigma \qquad \ldots(**)$$

Substituting in (*), we get

$$\therefore \qquad \sigma^2 \geq k^2\sigma^2 \left[ \int_{-\infty}^{\mu - k\sigma} f(x)\, dx + \int_{\mu + k\sigma}^{\infty} f(x)\, dx \right]$$

$$= k^2 \sigma^2 [P(X \leq \mu - k\sigma) + P(X \geq \mu + k\sigma)] \qquad \text{[From (**)]}$$

$$= k^2 \sigma^2 .\; P(|X - \mu| \geq k\sigma) \qquad \text{[From (**)]}$$

$$\Rightarrow \quad P(|X - \mu| \geq k\sigma) \leq 1/k^2, \qquad \ldots(***)$$

which establishes (6·73)

Also since

$P\{|X - \mu| \geq k\sigma\} + P\{|X - \mu| < k\sigma\} = 1$, we get

$P\{|X - \mu| < k\sigma\} = 1 - P\{|X - \mu| \geq k\sigma\} \geq 1 - \{1/k^2\}$ [From (***)]

which establishes (6·73*a*).

**Case (ii).** In case of discrete random variable, the proof follows exactly similarly on replacing integration by summation.

**Remark.** In particular, if we take $k\sigma = c > 0$, then (*) and (**) give respectively

$$P\left\{|X - \mu| \geq c\right| \leq \frac{\sigma^2}{c^2} \text{ and } P\left\{|X - \mu| < c\right\} \geq 1 - \frac{\sigma^2}{c^2}$$

$$\left.\begin{aligned} \Rightarrow \quad & P\left\{|X - E(X)| \geq c\right\} \leq \frac{\text{Var}(X)}{c^2} \\ \text{and} \quad & P\left\{|X - E(X)| < c\right\} \geq 1 - \frac{\text{Var}(X)}{c^2} \end{aligned}\right\} . \qquad \ldots(6\cdot73\, b)$$

**6·13·1. Generalised Form of Bienayme–Chebychev's Inequality.** *Let $g(X)$ be a non-negative function of a random variable X. Then for every $k > 0$, we have*

$$P\{ g(X) \geq k \} \leq \frac{E\{g(X)\}}{k} \qquad \ldots(6\cdot74)$$

**[Bangalore Univ. B.Sc., 1992]**

**Proof.** Here we shall prove the theorem for continuous random variable. The proof can be adapted to the case of discrete random variable on replacing integration by summation over the given range of the variable.

Let $S$ be the set of all $X$ where $g(X) \geq k$, *i.e.*,

$$S = \{x : g(x) \geq k\}$$

then $$\int_S dF(x) = P(X \in S) = P[g(X) \geq k], \qquad \ldots(*)$$

where $F(x)$ is the distribution function of $X$.

Now $$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, dF(x) \geq \int_S g(x)\, dF(x)$$

$$\geq k \,.\, P[g(X) \geq k] \quad [\because \text{ on } S,\ g(X) \geq k \text{ and using } (*)]$$

$$\Rightarrow \quad P[g(X) \geq k] \leq \frac{E[g(X)]}{k}$$

**Remarks 1.** If we take $g(X) = \{X - E(X)\}^2 = \{X - \mu\}^2$ and replace $k$ by $k^2 \sigma^2$ in (6·74), we get

$$P\left\{(X - \mu)^2 \geq k^2 \sigma^2\right\} \leq \frac{E(X - \mu)^2}{k^2 \sigma^2} = \frac{\sigma^2}{k^2 \sigma^2} = \frac{1}{k^2}$$

$$\Rightarrow \quad P\left\{|X - \mu| \geq k\sigma\right\} \leq 1/k^2, \qquad (6·74\,a)$$

which is Chebychev's inequality.

2. *Markov's Inequality.* Taking $g(X) = |X|$ in (6·74) we get, for any $k > 0$

$$P[|X| \geq k] \leq \frac{E|X|}{k}, \qquad \ldots(6·75)$$

which is Markov's inequality.

Rather, taking $g(X) = |X|^r$ and replacing $k$ by $k^r$ in (6·74), we get a more generalised form of Markov's inequality, *viz.*,

$$P[|X|^r \geq k^r] \leq \frac{E|X|^r}{k^r} \qquad \ldots(6·75a)$$

3. If we assume the existence of only second-order moments of $X$, then we cannot do better than Chebychev's inequality (6·73). However, we can sometimes improve upon the results of Chebychev's inequality if we assume the existence of higher order moments. We give below (without proof) one such inequality which assumes the existence of moments of 4th order.

**Theorem 6·31*a*.** $E|X|^4 < \infty$, $E(X) = 0$ *and* $E(X^2) = \sigma^2$

$$P\{|X| > k\sigma\} \geq \frac{\mu_4 - \sigma^4}{\mu_4 + \sigma^4 k^4 - 2k^2\sigma^4} \qquad \ldots(6·76)$$

If $X \sim U[0, 1]$, [*c.f.* Chapter 8], with p.d.f. $p(x) = 1,\ 0 < x < 1$ and $= 0$, otherwise, then

$$\begin{aligned} E(X^r) &= 1/(r+1);\ (r = 1, 2, 3, 4) \\ E(X) &= 1/2,\ E(X^2) = 1/3,\ E(X^3) = 1/4,\ E(X^4) = 1/5 \qquad \ldots(*) \\ \text{Var}(X) &= E(X^2) - [E(X)]^2 = 1/12 \\ \mu_4 &= E(X - \mu)^4 = E(X - \tfrac{1}{2})^4 = 1/80 \end{aligned}$$

[On using binomial expansion with (*)]

Chebychev's inequality (6·73a) with $k = 2$ gives :

$$P\left[\left|X - \tfrac{1}{2}\right| < 2\,\frac{1}{\sqrt{12}}\right] \geq 1 - \tfrac{1}{4} = 0{\cdot}75$$

With $k = 2$, (6·76) gives:

$$P\left[\left|X - \tfrac{1}{2}\right| > 2\,\frac{1}{\sqrt{12}}\right] \leq \frac{1/80 - 1/44}{1/80 + 1/9 - 1/8} = \frac{4}{49}$$

$$\Rightarrow \quad P\left[\left|X - \tfrac{1}{2}\right| \leq 2\,\frac{1}{\sqrt{12}}\right] \geq 1 - \frac{4}{49} = \frac{45}{49} = 0{\cdot}92,$$

which is a much better lower bound than the lower bound given by Chebychev's inequality.

**6·14. Convergence in probability.** We shall now introduce a new concept of convergenee, *viz.*, convergence in probability or stochastic convergence which is defined as follows:-

A sequence of random variables $X_1, X_2, ..., X_n, ...$ is said to converge in probability to a constant a, if for any $\varepsilon > 0$,

$$\lim_{n \to \infty} P(|X_n - a| < \varepsilon) = 1 \qquad ...(6{\cdot}77)$$

or its equivalent

$$\lim_{n \to \infty} P(|X_n - a| \geq \varepsilon) = 0 \qquad ...(6{\cdot}77a)$$

and we write

$$X_n \xrightarrow{p} a \text{ as } n \to \infty \qquad ...(6{\cdot}77b)$$

If there exists a random variable $X$ such that $X_n - X \xrightarrow{p} a$ as $n \to \infty$, then we say that the given sequence of random variables *converges in probability to the random variable X.*

**Remark. 1.** If a sequence of constants $a_n \to a$ as $n \to \infty$, then regarding the constant as a random variable having a one-point distribution at that point, we can say that as $a_n \xrightarrow{p} a$ as $n \to \infty$.

**2.** Although the concept of convergence in probability is basically different from that of ordinary convergence of sequence of numbers, it can be easily verified that the following simple rules hold for convergence in probability as well.

If $X_n \xrightarrow{p} \alpha$ and $Y_n \xrightarrow{p} \beta$ as $n \to \infty$, then

(*i*) $\quad X_n \pm Y_n \xrightarrow{p} \alpha \pm \beta$ as $n \to \infty$

(*ii*) $\quad X_n Y_n \xrightarrow{p} \alpha\beta$ as $n \to \infty$

(*iii*) $\dfrac{X_n}{Y_n} \xrightarrow{p} \dfrac{\alpha}{\beta}$ as $n \to \infty$, provided $\beta \neq 0$.

**6·14·1.** *(Chebychev's Theorem).* As an immediate consequence of Cheby -chev's inequality, we have the following theorem and convergence in probability.

"*If $X_1, X_2, ..., X_n$ is a sequence of random variables and if mean $\mu_n$ and standard deviation $\sigma_n$ of $X_n$ exists for all n and if $\sigma_n \to -$ as $n \to \infty$, then*

$$X_n - \mu_n \xrightarrow{p} 0 \text{ as } n \to \infty$$

**Proof.** We know, for any $\varepsilon > 0$

$$P\left\{ \left| X_n - \mu_n \right| \geq \varepsilon \right\} \leq \frac{\sigma_n^2}{\varepsilon^2} \to 0 \text{ as } n \to \infty$$

Hence $X_n - \mu_n \xrightarrow{p} 0$ as $n \to \infty$

**6·15. Weak Law of Large Numbers.** *Let $X_1, X_2, ..., X_n$ be a sequence of random variables and $\mu_1, \mu_2, ..., \mu_n$ be their respective expectations and let*

$$B_n = \text{Var}(X_1 + X_2 + \ldots + X_n) < \infty$$

$$\text{Then } P\left\{ \left| \frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{\mu_1 + \mu_2 + \ldots + \mu_n}{n} \right| < \varepsilon \right\} \geq 1 - \eta \qquad \ldots(6\cdot78)$$

*for all $n > n_0$, where $\varepsilon$ and $\eta$ are arbitrary small positive numbers, provided*

$$\lim_{n \to \infty} \frac{B_n}{n^2} \to 0$$

**Proof.** Using Chebychev's Inequality (6·73*b*), to the random variable $(X_1 + X_2 + \ldots + X_n)/n$, we get for any $\varepsilon > 0$,

$$P\left\{ \left| \left( \frac{X_1 + X_2 + \ldots + X_n}{n} \right) - \left( E\,\frac{X_1 + X_2 + \ldots + X_n}{n} \right) \right| < \varepsilon \right\} \geq 1 - \frac{B_n}{n^2 \varepsilon^2},$$

$$\left[ \text{since Var}\left( \frac{X_1, X_2 + \ldots + X_n}{n} \right) = \frac{1}{n^2}\,\text{Var}(X_1 + X_2 + \ldots + X_n) = \frac{B_n}{n^2} \right]$$

$$\Rightarrow \quad P\left\{ \left| \frac{X_1, X_2 + \ldots + X_n}{n} - \frac{\mu_1 + \mu_2 + \ldots + \mu_n}{n} \right| < \varepsilon \right\} \geq 1 - \frac{B_n}{n^2 \varepsilon^2}.$$

So far, nothing is assumed about the behaviour of $B_n$ for indefinitely increasing values of $n$. Since $\varepsilon$ is arbitrary, we assume $\dfrac{B_n}{n^2 \varepsilon^2} \to 0$, as $n$ becomes indefinitely large. Thus, having chosen two arbitary small positive numbers $\varepsilon$ and $\eta$, number $n_0$ can be found so that the inequality

$$\frac{B_n}{n^2 \varepsilon^2} < \eta,$$

will hold for $n > n_0$. Consequently, we shall have

$$P\left\{ \left| \frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{\mu_1 + \mu_2 + \ldots + \mu_n}{n} \right| < \varepsilon \right\} \geq 1 - \eta$$

for all $n > n_0(\varepsilon, \eta)$.

This conclusion leads to the following important result, known as the (*Weak*) Law of Large Numbers:

*"With the probability approaching unity or certainty as near as we please, we may expect that the arithmetic mean of values actually assumed by n random variables will differ from the arithmetic mean of their expectations by less than any given number, however small, provided the number of variables can be taken sufficently large and provided the condition*

$$\frac{B_n}{n^2} \to 0 \text{ as } n \to \infty$$

*is fulfilled".*

**Remarks. 1.** Weak law of large numbers can also be stated as follows:

$$\bar{X}_n \xrightarrow{p} \bar{\mu}_n$$

provided $\frac{B_n}{n^2} \to 0$ as $n \to \infty$, symbols having their usual meanings.

**2.** For the existence of the law we assume the following conditions:

(*i*) $E(X_i)$ exists for all $i$, (*ii*) $B_n = \text{Var}(X_1 + X_2 + \ldots + X_n)$ exists, and

(*iii*) $B_n/n^2 \to 0$ as $n \to \infty$.

Condition (*i*) is necessary, without it the law itself cannot be stated. But the conditions (*ii*) and (*iii*) are not necessary, (*iii*) is however a sufficient condition.

**3.** If the variables $X_1, X_n, \ldots, X_n$ are independent and identically distributed, *i.e.*, if $E(X_i) = \mu$ (say), and $\text{Var}(X_i) = \sigma^2$ (say) for all $i = 1, 2, \ldots, n$ then

$$B_n = \text{Var}(X_1 + X_2 + \ldots + X_n) = \sum_{i=1}^{n} \text{Var}(X_i)$$

the convariance terms vanish, since variables are independent.

$$\therefore \qquad B_n = n\,\sigma^2 \qquad \ldots(*)$$

Hence

$$\lim_{n \to \infty} \frac{B_n}{n^2} = \lim_{n \to \infty} (\sigma^2/n) = 0$$

Thus, the law of large number holds for the sequence $\{X_n\}$ of *i.i.d. r.v.'s* and we get

$$P\left\{ \left| \frac{X_1 + X_2 + \ldots + X_n}{n} - \mu \right| < \varepsilon \right\} > 1 - \eta \ \forall\ n > n_0$$

*i.e.*, $\quad P\{|\bar{X}_n - \mu| < \varepsilon\} \to 1$ as $n \to \infty$

$\Rightarrow \quad P\{|\bar{X}_n - \mu| \geq \varepsilon\} \to 0$ as $n \to \infty$

where $\bar{X}_n$ is the mean of the $n$ random variables $X_1, X_2, \ldots, X_n$. This result implies that $\bar{X}_n$ converges in *probability* to $\mu$, *i.e.*,

$$\bar{X}_n \xrightarrow{p} \mu$$

**Note.** If $\overline{X}_n$ is the mean of $n$ *i.i.d.* r.v.'s $X_1, X_2, \ldots, X_n$ with

$E(X_i) = \mu$ ; Var $(X_i) = \sigma^2$, then [On using (*)]

$$E(\overline{X}_n) = \mu \text{ ; and } \text{Var}(\overline{X}_n) = \text{Var}\left(\sum_{i=1}^{n} x_i / n\right) \qquad \ldots(6\cdot80)$$

**Theorem 6·32.** *If the variables are uniformly bounded then the condition,*

$$\lim_{n \to \infty} \frac{B_n}{n^2} = 0$$

*is necessary as well as sufficient for WLLN to hold.*

**Proof.** Let $\xi_i = X_i - a_i$, where $E(X_i) = a_i$ ; then $E(\xi_i) = 0$, $(i = 1, 2, \ldots, n)$.

Since $X_i$'s are uniformly bounded, there exists a positive number $c < \infty$ such that $|\xi_i| < c$.

If $\quad p = P[\,|\xi_1 + \xi_2 + \ldots + \xi_n| \le n\,\varepsilon]$

then $\quad 1 - p = P[\,|\xi_1 + \xi_2 + \ldots + \xi_n| > n\,\varepsilon]$

Let $\quad U_n = \xi_1 + \xi_2 + \ldots + \xi_n$,

then $\quad E(U_n) = \sum_{i=1}^{n} E(\xi_i) = 0$

and $\quad \text{Var}(U_n) = E(U_n^2) = B_n$ (say).

$$\Rightarrow \quad B_n = \int_0^\infty U_n^2\, dF(U_n), \text{ where } F(U_n) \text{ is d.f. of } U_n.$$

$$= \int_{U_n^2 \le n^2 \varepsilon^2} U_n^2\, dF + \int_{U_n^2 > n^2 \varepsilon^2} U_n^2\, dF$$

$$\le n^2 \varepsilon^2 \int_{|U_n| \le n\varepsilon} dF + n^2 c^2 \int_{|U_n| > n\varepsilon} dF$$

$$\le n^2 \varepsilon^2 p + n^2 c^2 (1 - p)$$

$$\therefore \quad \frac{B_n}{n^2} \le \varepsilon^2 p + c^2 (1 - p)$$

If the law of large numbers holds,

$$1 - p = P[\,|\xi_1 + \xi_2 + \ldots + \xi_n| > n\,\varepsilon\,] \to 0 \text{ as } n \to \infty.$$

Hence as $n \to \infty$, $(1 - p) \to 0$, and

$\frac{B_n}{n^2} < \varepsilon^2 p + c^2 \delta$, $\varepsilon$ and $\delta$ being arbitrarily small positive numbers.

Hence $\quad \frac{B_n}{n^2} \to 0$ as $n \to \infty$.

**6·15·1. Bernoulli's Law of Large Numbers.** Let there be $n$ trials of an event, each trial resulting in a success or failure. If $X$ is the number of successes in $n$ trials with constant probability $p$ of success for each trial, then $E(X) = np$ and Var $(X) = npq$, $q = 1 - p$. The variable $X/n$ represents the proportion of successes or the relative frequency of successes, and

$$E(X/n) = \frac{1}{n} E(X) = p, \text{ and } \operatorname{Var}(X/n) = \frac{1}{n^2} \operatorname{Var}(X) = \frac{pq}{n}$$

Then

$$P\left\{ \left| \frac{X}{n} - p \right| < \varepsilon \right\} \to 1 \text{ as } n \to \infty \qquad \text{...(6·81)}$$

$$\Rightarrow \quad P\left\{ \left| \frac{X}{n} - p \right| \geq \varepsilon \right\} \to 0 \text{ as } n \to \infty \qquad \text{...(6·81 } a)$$

tor any assigned $\varepsilon > 0$. This implied that $(X/n)$ converges in probability to $p$ as $n \to \infty$.

**Proof.** Applying Chebychev's Inequality [Form (6·73 *b*)] to the variable $X/n$, we get for any $\varepsilon > 0$;

$$P\left\{ \left| \frac{X}{n} - E\left(\frac{X}{n}\right) \right| \geq \varepsilon \right\} \frac{\operatorname{Var}(X/n)}{\varepsilon^2}$$

$$\Rightarrow \quad P\left\{ \left| \frac{X}{n} - p \right| < \varepsilon \right\} \to 1 \text{ as } n \to \infty$$

since the maximum value of $pq$ is at $p = q = 1/2$ *i.e.*, max $(p\, q) = 1/4$ *i.e.*, $pq \leq 1/4$.

Since $\varepsilon$ is arbitrary, we get

$$P\left\{ \left| \frac{X}{n} - p \right| \geq \varepsilon \right\} \to 0 \text{ as } n \to \infty$$

$$\Rightarrow \quad P\left\{ \left| \frac{X}{n} - p \right| < \varepsilon \right\} \to 1 \text{ as } n \to \infty$$

**6·15·2. MorkofPs Theorem.** *The law of large numbers holds if for some $\delta > 0$, all the mathematical expectations*

$$E(|X_i|^{1+\delta}) \,;\, i = 1, 2, \ldots \qquad \text{...(6·82)}$$

*exist and are bounded.*

**6·15·3. Khinchin's Theorem.** *If $X_i$'s are identically and independently distributed random variables, the only condition necessary for the law of large numbers to hold is that $E(X_i)$ ; $i = 1, 2, \ldots$ should exist.*

**Theorem 6·33.** *Let $\{X_n\}$ be any sequence of r.v.'s. Write :*

$$Y_n = [\, S_n - E(S_n)\,]/n \text{ where } S_n = X_1 + X_2 + \ldots + X_n .$$

*A necessary and sufficient condition for the sequence $\{X_n\}$ to satisfy the W.L.L.N. is that*

$$E\left\{ \frac{Y_n^2}{1 + Y_n^2} \right\} \to 0 \text{ as } n \to \infty . \qquad \text{...(6·83)}$$

**Proof.** *If Part:* Let us assume that (6·83) holds. We shall prove $\{X_n\}$ satisfies W.L.L.N.

For real numbers $a, b$ ; $a \geq b > 0$ we have:

$$a \geq b \quad \Rightarrow \quad a + ab \geq b + ab \qquad \ldots(*)$$

Let us define the event $A = \{ |Y_n| \geq \varepsilon \}$.

$$w \in A_X \quad \Rightarrow \quad |Y_n| \geq \varepsilon \quad \Rightarrow \quad |Y_n|^2 \geq \varepsilon^2 > 0$$

$\therefore$ Taking $a = Y_n^2$ and $b = \varepsilon^2$ in (*), we define another event $B$ as follows:

$$B_n = \left\{ \left( \frac{Y^{n^2}}{1 + Y_n^2} \right) \left( \frac{1 + \varepsilon^2}{\varepsilon^2} \right) \geq 1 \right\} = \left\{ \frac{Y_n^2}{1 + Y_n^2} \geq \frac{\varepsilon^2}{1 + \varepsilon^2} \right\}$$

Since $w \in A \quad \Rightarrow \quad w \in B, \quad A \subseteq B \quad \Rightarrow \quad P(A) \leq P(B)$

$$\therefore \qquad P\left[ |Y_n| \geq \varepsilon \right] \leq P\left[ \frac{y_n^2}{1 + Y_n^2} \frac{\varepsilon^2}{1 + \varepsilon^2} \right]$$

$$\leq \frac{E\{ Y_n/(1 + Y_n^2) \}}{\varepsilon^2/(1 + \varepsilon^2)}$$

[By Markov's Inequality (6·75)]

$\to 0$ as $n \to \infty$ [By assumption (6·83)]

$$\therefore \qquad P\left[ |Y_n| \geq \varepsilon \right] \to 0 \text{ as } n \to \infty$$

$$\Rightarrow \qquad \lim_{n \to \infty} P\left[ \left| \frac{S_n - E(S_n)}{n} \right| \geq \varepsilon \right] \to 0$$

$\Rightarrow$ WLLN holds for the sequence $\{X_n\}$ of r.v.'s.

Conversely, if $\{X_n\}$ satisfies WLLN, we shall establish (6·83). Let us assume that $X_i$'s are continuous and let $Y_n$ have p.d.f. $f_n(y)$. Then

$$E\left\{ \frac{Y_n^2}{1 + Y_n^2} \right\} = \int_{-\infty}^{\infty} \frac{y^2}{1 + y^2} \cdot f_n(y)\, dy$$

$$= \left( \int_A + \int_{A^c} \right) \frac{y^2}{1 + y^2} f_n(y)\, dy$$

where $\qquad A = \{ |Y| \geq \varepsilon \}$ and $A^c = \{ |y| < \varepsilon \}$

$$E\left(\frac{Y_n^2}{1+Y_n^2}\right) \le \int_A 1 . f_n(y)\, dy + \int_{A^c} y^2 . f_n(y)\, dy \;\Bigg|\; \because \frac{y^2}{1+y^2} < 1 \text{ and } \frac{y^2}{1+y^2} < y^2 \Bigg|$$

$$\le P(A) + \varepsilon^2 \int_{A^c} f_n(y)\, dy \qquad (\because \text{On } A^c : |y| < \varepsilon)$$

$$= P(A) + \varepsilon^2 . P(A^c)$$

$$\le P(A) + \varepsilon^2 \qquad (\because P(A^c) < 1)$$

$$\Rightarrow E\left[\frac{Y_n^2}{1+Y_n^2}\right] \le P\{|Y_n| \ge \varepsilon\} + \varepsilon^2 \qquad \ldots(**)$$

But since $\{X_n\}$ satisfies WLLN, we have :

$$\lim_{n \to \infty} P[|Y_n| \ge \varepsilon] \to 0$$

and since $\varepsilon^2$ is arbitrarily small positive number, we get on taking limits in (**).

$$\lim_{n \to \infty} E\left[\frac{Y_n^2}{1+Y_n^2}\right] \to 0$$

**Corollary.** Let $X_1, X_2, \ldots, X_n$ be sequence of independent r.v.'s such that Var $(X_i) < \infty$ for $i = 1, 2, \ldots$ and

$$\frac{B_n}{n^2} = \frac{\text{Var}\left(\sum_{i=1}^{n} X_i\right)}{n^2} = \frac{\text{Var}(S_n)}{n^2} \to 0 \text{ as } n \to \infty$$

Then WLLN holds.

**Proof.** We have :

$$\frac{Y_n^2}{1+Y_n^2} \le Y_n^2 = \left[\frac{S_n - E(S_n)}{n}\right]^2$$

$$\Rightarrow \qquad E\left[\frac{Y_n^2}{1+Y_n^2}\right] \le \frac{1}{n^2} \cdot E[S_n - E(S_n)]^2 = \frac{\text{Var } S_n}{n^2} = \frac{B_n}{n^2}$$

$$\therefore \qquad \lim_{n \to \infty} E\left[\frac{Y_n^2}{1+Y_n^2}\right] \le \lim_{n \to \infty} \frac{B_n}{n^2} \to 0 \qquad \text{(By assumption)}$$

Hence by the above theorem WLLN holds for the sequence $\{X_n\}$ of r.v.'s.

**Remark.** The result of Theorem 6·33 holds even if $E(X_i)$ does not exist. In this case we simply define $Y_n = [S_n/n]$ rather than $[S_n - E(S_n)]/n$.

**Example 6·48.** *A symmetric die is thrown 600 times. Find the lower bound for the probability of getting 80 to 120 sixes.*

**Solution.** Let $S$ be total number of successes.

Then

$$E(S) = np = 600 \times \frac{1}{6} = 100$$

$$V(S) = npq = 600 \times \frac{1}{6} \times \frac{5}{6} = \frac{500}{6}$$

Using Chebychev's inequality, we get

$$P[\,|S - E(S)| < k\sigma\,] \geq 1 - \frac{1}{k^2}$$

$$\Rightarrow \qquad P\{\,|S - 100| < k\sqrt{500/6}\,\} \geq 1 - \frac{1}{k^2}$$

$$\Rightarrow \qquad P\{\,100 - k\sqrt{500/6} < S < 100 + k\sqrt{500/6}\,\} \geq 1 - \frac{1}{k^2}$$

Taking $\quad k = \dfrac{20}{\sqrt{500/6}}$, we get

$$P(80 \leq S \leq 120) \geq 1 - \frac{1}{400 \times (6/500)} = \frac{19}{24}$$

**Example 6·49.** *Use Chebychev's inequality to determine how many times a fair coin must be tossed in order that the probability will be at least 0·90 that the ratio of the observed number of heads to the number of tosses will lie between 0·4 and 0·6.*

**[Madras Univ. B.Sc (Stat.)Oct 1991; Delhi Univ. B.Sc. (Stat Hons.) 1989]**

**Solution.** As in the proof of Bernoulli's Law of Large Numbers, we get for any $\varepsilon > 0$,

$$P\left\{\left|\frac{X}{n} - p\right| \geq \varepsilon\right\} \leq \frac{1}{4n\,\varepsilon^2}$$

$$\Rightarrow \qquad P\left\{\left|\frac{X}{n} - p\right| < \varepsilon\right\} \geq 1 - \frac{1}{4n\,\varepsilon^2}$$

Since $p = 0{\cdot}5$ (as the coin is unbaised) and we want the proportion of successes $X/n$ to lie between 0·4 and 0·6, we have

$$\left|\frac{X}{n} - p\right| \leq 0{\cdot}1$$

Thus choosing $\varepsilon = 0{\cdot}1$, we have

$$P\left\{\left|\frac{X}{n} - p\right| < 0{\cdot}1\right\} \geq 1 - \frac{1}{4n\,(0{\cdot}1)^2} = 1 - \frac{1}{0{\cdot}04\,n}$$

Since we want this probability to be 0·9, we fix

$$1 - \frac{1}{0{\cdot}04\,n} = 0{\cdot}90$$

$$\Rightarrow \qquad 0{\cdot}10 = \frac{1}{0{\cdot}04n}$$

$$\Rightarrow \qquad n = \frac{1}{0{\cdot}10 \times 0{\cdot}04} = 250$$

Hence the required number of tosses is 250.

**Example 6·50.** *For geometric distribution* $p(x) = 2^{-x}$ ; $x = 1, 2, 3, \ldots$ *prove that Chebychev's inequality gives*

$$P[\,|X-2| \le \,] > \tfrac{1}{2}$$

*while the actual probability is* $\frac{15}{16}$. **[Nagpur Univ. B.Sc. (Stat.), 1989]**

**Solution.** $E(X) = \sum_{x=1}^{\infty} \frac{x}{2^x} = \frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \frac{4}{2^4} + \ldots$

$$= \frac{1}{2}(1 + 2A + 3A^2 + 4A^3 + \ldots), \quad (A = 1/2)$$

$$= \frac{1}{2}(1-A)^{-2} = 2$$

$$E(X^2) = \sum_{x=1}^{\infty} \frac{x^2}{2^x} = \frac{1}{2^2} + \frac{4}{2^3} + \frac{9}{2^4} + \ldots$$

$$= \frac{1}{4}[1 + 4A + 9A^2 + \ldots] = \frac{1}{4}(1+A)(1-A)^{-3} = 6$$

[See Example 6·17]

$\therefore \qquad \text{Var}(X) = E(X^2) - [E(X)]^2 = 6 - 4 = 2$

Using Chebychev's inequality, we get

$$P\{|X - E(X)| \le k\sigma\} > 1 - \frac{1}{k^2}$$

With $k = \sqrt{2}$, we get

$$P\{|X-2| \le \sqrt{2}\,.\sqrt{2}\} > 1 - \tfrac{1}{2} = \tfrac{1}{2}$$

$$\Rightarrow \qquad P\{|X-2| \le 2\} > \tfrac{1}{2}$$

And the actual probability is given by

$$P\{|X-2| \le 2\} = P\{0 \le X \le 4\} = P\{X = 1, 2, 3 \text{ or } 4\}$$

$$= \tfrac{1}{2} + (\tfrac{1}{2})^2 + (\tfrac{1}{2})^3 + (\tfrac{1}{2})^4 = \frac{15}{16}$$

**Example 6·51.** *Does there exist a variate X for which*

$$P[\mu_x - 2\sigma \le X \le \mu_x + 2\sigma] = 0{\cdot}6 \qquad \ldots(*)$$

**[Delhi Univ. B.Sc.(Maths Hons.) 1983]**

**Solution.** We have :

$$P[\mu_x - 2\sigma \le X \le \mu_x + 2\sigma] = P[|X - \mu_x| \le 2\sigma]$$

$$\ge 1 - \frac{1}{4} = 0{\cdot}75 \qquad \text{(Using Chebychev's Inequality)}$$

Since lower bound for the probability is 0·75, there does not exist a r.v. $X$ for which (*) holds.

**Example 6·52.** *(a) For the discrete variate with density*

$$f(x) = \frac{1}{8} I_{(-1)}(x) + \frac{6}{8} I_{(0)}(x) + \frac{1}{8} I_{(1)}(x)$$

*evaluate* $P[|X - \mu_x| \geq 2\sigma_x]$. **[Delhi Univ. B.Sc.(Maths Hons.) 1989]**

*(b)* Compare this result with that obtained on using Chebychev's inequality.

**Hint.** *(a)* Here $X$ has the probability distribution :

| | | | | |
|---|---|---|---|---|
| $x$ : | $-1$ | $0$ | $1$ | $\therefore\ E(X) = -1 \times \frac{1}{8} + 1 \times \frac{1}{8} = 0$ |
| $p(x)$ : | 1/8 | 6/8 | 1/8 | $EX^2 = 1 \times \frac{1}{8} + 1 \times \frac{1}{8} = \frac{1}{4}$ |

$$\therefore \quad \text{Var}(X) = E(X^2) - [E(X)]^2 = 1/4 \quad \Rightarrow \quad \sigma_x = 1/2$$

$$P[|X - \mu_x| \geq 2\sigma_x] = P[|X| \geq 1] = 1 - P(|X| < 1)$$
$$= 1 - P[-1 < X < 1] = 1 - P(X = 0) = 1/4$$

*(b)* $$P[|X - \mu_x| \geq 2\sigma_x] \leq \frac{1}{4} \qquad \text{(By Chebychev's Inequality)}$$

In this case both results are same.

**Note.** This example shows that, in general, Chebychev's inequality cannot be improved.

**Example 6·53.** *Two unbiased dice are thrown. If $X$ is the sum of the numbers showing up, prove that*

$$P(|X - 7| \geq 3) \leq \frac{35}{54}.$$

*Compare this with the actual probability.* **(Karnataka Univ. B.Sc., 1988)**

**Solution.** The probability distribution of the r.v. $X$ (the sum of the numbrs on the two dice) is as given below :

| $X$ | Favourable cases (distinct) | Probability ($p$) |
|---|---|---|
| 2 | (1, 1) | 1/36 |
| 3 | (1, 2), (2, 1) | 2/36 |
| 4 | (1, 3), (3, 1), (2, 2) | 3/36 |
| 5 | (1, 4), (4, 1), (2, 3), (3, 2) | 4/36 |
| 6 | (1, 5), (5, 1), (2, 4), (4, 2), (3, 3) | 5/36 |
| 7 | (1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3) | 6/36 |
| 8 | (2, 6), (6, 2), (3, 5), (5, 3), (4, 4) | 5/36 |
| 9 | (3, 6), (6, 3), (4, 5), (5, 4) | 4/36 |
| 10 | (4, 6), (6, 4), (5, 5) | 3/36 |
| 11 | (5, 6), (6, 5) | 2/36 |
| 12 | (6, 6) | 1/36 |

$$E(X) = \sum_x p \cdot x$$
$$= \frac{1}{36}(2 + 6 + 12 + 20 + 30 + 42 + 40 + 36 + 30 + 22 + 12)$$
$$= \frac{1}{36}(252) = 7$$
$$E(X^2) = \sum_x p \cdot x^2$$
$$= \frac{1}{36}[4 + 18 + 48 + 100 + 180 + 294 + 320 + 324 + 300 + 242 + 144]$$
$$= \frac{1}{36}(1974) = \frac{1974}{36} = \frac{329}{6}$$
$$\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{329}{6} - (7)^2 = \frac{35}{6}$$

By Chebychev's inequality, for $k > 0$, we have

$$P(|X - \mu| \geq k) \leq \frac{\text{Var}\,X}{k^2}$$
$$\Rightarrow \quad P(|X - 7| \geq 3) \leq \frac{35/6}{9} = \frac{35}{54} \qquad \text{(Taking } k = 3)$$

**Actual Probability :**

$$P(|X - 7| \geq 3) = 1 - P(|X - 7| < 3$$
$$= 1 - P(4 < X < 10)$$
$$= 1 - [P(X = 5) + P(X = 6) + P(X = 7)$$
$$P(X = 8) + P(X = 9)]$$
$$= 1 - \frac{1}{36}[4 \quad 5 + 6 + 5 + 4] = 1 - \frac{24}{36} = \frac{1}{3}$$

**Example 6·54.** *If $X$ is the number scored in a throw of a fair die, show that the Chebychev's inequality gives*

$$P[|X - \mu| > 2{\cdot}5] < 0{\cdot}47,$$

*where $\mu$ is the mean of $X$, while the actual probability is zero.*

**[Kerala Univ. B.Sc., Oct. 1989]**

**Solution.** Here $X$ is a random variable which takes the values 1,2,...,6, each with probability 1/6. Hence

$$E(X) = \tfrac{1}{6}(1 + 2 + \ldots + 6) = \tfrac{1}{6} \cdot \frac{6 \times 7}{2} = \frac{7}{2}$$
$$E(X^2) = \tfrac{1}{6}(1^2 + 2^2 + \ldots + 6^2) = \tfrac{1}{6} \frac{6 \times 7 \times 13}{6} = \frac{91}{6}$$
$$\therefore \quad \text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12} = 2{\cdot}9167$$

For $k > 0$, Chebychev's inequality gives

$$P[\,|X-E(X)|>k\,]<\frac{\text{Var}\,X}{k^2}$$

Choosing $k = 2{\cdot}5$, we get

$$P[\,|X-\mu|>2{\cdot}5\,]<\frac{2{\cdot}9167}{6{\cdot}25}=0{\cdot}47$$

The actual probability is given by

$$p = P[\,|X-3{\cdot}5|>2{\cdot}5\,]$$

$$= P[\,X \text{ lies outside the limits } (3{\cdot}5-2{\cdot}5,\ 3{\cdot}5+2{\cdot}5),\ i.e.,\ (1,6)]$$

But since $X$ is the numberon the dice when thrown, it cannot lie outside the limits of 1 and 6.

$$\therefore \quad p = p(\varphi) = 0$$

**Example 6·55.** *If the variable $X_p$ assumes the value $2^{p-2\log p}$ with probability $2^{-p}$; $p = 1, 2, \ldots$, examine if the law of large numbers holds in this case.*

**Solution.** Putting $p = 1, 2, 3, \ldots$, we get

$$2^{1-2\log 1},\ 2^{2-2\log 2},\ 2^{3-2\log 3},\ \ldots$$

as the values of the variables $X_1, X_2, X_3 \ldots$respectively, and

$$\frac{1}{2},\ \frac{1}{2^2},\ \frac{1}{2^3},\ \ldots$$

their corresponding probabilities. Therefore,

$$E(X_k) = 2^{1-2\log 1}\cdot\frac{1}{2}+2^{2-2\log 2}\cdot\frac{1}{2^2}+\ldots=\sum_{p=1}^{\infty}2^{p-2\log p}\cdot 2^{-p}$$

$$=\sum_{p=1}^{\infty}\frac{1}{2^{2\log p}}$$

Let $\quad U = 2^{2\log p}$, then

$$\log U = 2\log p\log 2 = \log p\,.\,\log 2^2 = \log 4\,.\,\log p = \log p^{\log 4}$$

$$\therefore \quad E(X_k)=\sum_{p=1}^{\infty}\frac{1}{p^{\log 4}} = 1+\frac{1}{2^{\log 4}}+\frac{1}{3^{\log 4}}+\ldots$$

which is a convergent series.

$$\left[\because \sum_{n=1}^{\infty}\frac{1}{n^p} \text{ is convergent if and only if } p>1\right]$$

Therefore, the mathematical expectation of the variables $X_1, X_2, \ldots$, exists. Thus by Khinchin's theorem, the law of large numbers holds in this case.

**Exmaple 6·56.** *Let $X_1, X_2, \ldots, X_n$ be i.i.d. variables with mean $\mu$ and variance $\sigma^2$ and as $n \to \infty$,*

$$(X_1^2+X_2^2+\ldots+X_n^2)/n \xrightarrow{p} c,$$

*for some constant $c$; $(0 \le c \le \infty)$. Find $c$.* **[Delhi Univ. B.Sc. (Stat. Hons.), 1989]**

**Solution.** We are given $E(X_i) = \mu$ ; $\text{Var}(X_i) = \sigma^2$ ; $i = 1, 2, ..., n$ .

$\therefore$ $E(X_i^2) = \text{Var}(X_i) + [E(X_i)]^2 = \sigma^2 + \mu^2$ (finite) ; $i = 1, 2, ..., n$ .

Since $E(X_i^2)$ is finite; by Khinchines's Theorem WLLN holds for the sequence $\{X_i^2\}$ of i.i.d. r.v.'s so that

$$(X_1^2 + X_2^2 + \dots + X_n^2)/n \xrightarrow{p} E(X_i^2), \text{ as } n \to \infty$$

$$\Rightarrow \quad (X_1^2 + X_2^2 + \dots + X_n^2)/n \xrightarrow{p} \sigma^2 + \mu^2 = c, \text{ as } n \to \infty$$

Hence $\quad c = \sigma^2 + \mu^2$ .

**Example 6·57.** *How large a sample must be taken in order that the probability will be at least 0·95 that $\overline{X}_n$ will be lie within 0·5 of $\mu$ . $\mu$ is unknown and $\sigma = 1$ .* **[Delhi Univ. B.Sc. (Maths Hons.) 1988]**

**Solution.** We have: $E(\overline{X}_n) = \mu$ and $\text{Var}(\overline{X}_n) = \sigma^2/n$

[ c.f. § 6·15, = $n$ (6·80)]

Applying Chebychev's inequality to the r.v. $\overline{X}_n$ we get, for any $c > 0$

$$P[|\overline{X}_n - E(\overline{X}_n)| < c] \geq 1 - \frac{\text{Var}(\overline{X}_n)}{c^2}$$

$$\Rightarrow \quad P[|\overline{X}_n - \mu| < c] \geq 1 - \frac{\sigma^2}{n c^2} \quad \quad ...(*)$$

We want $n$ so that

$$P[|\overline{X}_n - \mu| < 0{\cdot}5] \geq 0{\cdot}95 \quad \quad ...(**)$$

Comparing (*) and (**) we get :

$$c = 0{\cdot}5 = 1/2 \quad \text{and} \quad 1 - \frac{\sigma^2}{n c^2} = 0{\cdot}95 \quad \text{and} \quad \sigma = 1 \quad \text{(Given)}$$

$$\therefore \quad 1 - \frac{4}{n} = 0{\cdot}95 \Rightarrow \frac{4}{n} = 0{\cdot}05 = \frac{1}{20} \Rightarrow n = 80.$$

Hence $n \geq 80$ .

**Example 6·58.** *(a) Let $X_i$ assume that values $i$ and $-i$ with equal probabilities. Show that the law of large numbers cannot be applied to the independent variables $X_1, X_2, ...,$ i.e., $X_i$'s.*

*(b) If $X_i$ can have only two values with equal probabilities $i^\alpha$ and $-i^\alpha$, show that the law of large numbers can be applied to the independent variables $X_1, X_2, ...,$ if $\alpha < \frac{1}{2}$.*

**Solution.** (*a*) We have

$$P(X_i = i) = \tfrac{1}{2}, \quad P(X_i = -i) = \tfrac{1}{2}$$

$$E(X_i) = \tfrac{1}{2}(i) + \tfrac{1}{2}(-i) = 0 \; ; \; i = 1, 2, 3, \dots$$

$$V(X_i) = E(X_i^2) = \frac{i^2}{2} + \frac{i^2}{2} = i^2 \quad [\because E(X_i) = 0] \quad \quad ...(*)$$

$$B_n = V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$$

$$= (1 + 2^2 + \dots + n^2) = \frac{n(n+1)(2n+1)}{6} \qquad \text{...[From (*)]}$$

$\therefore \quad \frac{B_n}{n^2} \to \infty$ as $n \to \infty$. Hence we cannot draw any conclusion whether WLLN holds or not. Here, we need to apply further tests. (See Theorem 6·33 page 6·104)

(b) $E(X_i) = \frac{i^{\alpha}}{2} + \left(\frac{-i^{\alpha}}{2}\right) = 0$

$$E(X_i^2) = \frac{(i^{\alpha})^2}{2} + \frac{(-i^{\alpha})^2}{2} = i^{\alpha 2}$$

$$V(X_i) = E(X_i^2) - [E(X_i)]^2 = i^{2\alpha}$$

$$B_n = V(X_1 + X_2 + \dots + X_n) = \sum_{i=1}^{n} V(X_i) = \sum_{i=1}^{n} i^{2\alpha}$$

$$= 1^{2\alpha} + 2^{2\alpha} + \dots + 2^{2\alpha} = \int_0^n x^{2a}\, dx$$

[From Euler Maclaurin's Formula]

$$= \left| \frac{x^{2\alpha+1}}{2\alpha+1} \right|_0^n = \frac{n^{2\alpha+1}}{2\alpha+1}$$

$$\therefore \quad \frac{B_n}{n^2} = \frac{n^{2\alpha-1}}{2\alpha+1} \to 0 \text{ if } 2\alpha - 1 < 0 \Rightarrow \alpha < \tfrac{1}{2}$$

Hence the result.

**Example 6·59.** Let $\{X_k\}$ be mutually independent and identically distributed random variables with mean $\mu$ and finite variance. If $S_n = X_1 + X_2 + \dots + X_n$, prove that the law of large numbers does not hold for the sequence $\{S_n\}$.

**Solution.** The variables now are $S_1, S_2, \dots, S_n$.

$$B_n = V(S_1 + X_2 + \dots + S_n)$$

$$= V\{X_1 + (X_1 + X_2) + (X_1 + X_2 + X_3) + \dots + (X_1 + X_2 + \dots + X_n)\}$$

$$= V\{nX_1 + (n-1)X_2 + \dots + 2X_{n-1} + X_n\}$$

$$= n^2 V(X_1) + (n-1)^2 V(X_2) + \dots + 2^2 V(X_{n-1}) + 1^2 V(X_n)$$

(Covariance terms vanish since variables are independent.)

Let $V(X_i) = \sigma^2$ for all $i$.

(Since the variables are identically distributed.)

$$\therefore \quad B_n = (1^2 + 2^2 + \dots + n^2)\sigma^2 + \frac{n(n+1)(2n+1)}{6}\sigma^2$$

$$\therefore \quad \frac{B_n}{n^2} = \frac{(n+1)(2n+1)}{6n} \cdot \sigma^2 \to \infty \text{ as } n \to \infty$$

**Example 6·60.** *Examine whether the week law of large numbers holds for the sequence* $\{X_k\}$ *of independent random variables defined as follows:*

$$P[X_k = \pm 2^k] = 2^{-(2k+1)}$$

$$P[X_k = 0] = 1 - 2^{-2k}$$ **[Delhi Univ. B.Sc. (Maths Hons.), 1988]**

**Solution.** We have

$$E(X_k) = 2^k\, 2^{-(2k+1)} + (-2^k) \cdot 2^{-(2k+1)} + 0 \times (1 - 2^{-2k})$$

$$= 2^{-(2k+1)}\,(2^k - 2^k) = 0$$

$$E(X_k^2) = (2^k)^2 \cdot 2^{-(2k+1)} + (-2^k)^2 \cdot 2^{-(2k+1)} + 0^2 \times (1 - 2^{-2k})$$

$$= 2^{2k} \cdot 2^{-(2k+1)} + 2^{k^2} \cdot 2^{-(2k+1)}$$

$$= \tfrac{1}{2} + \tfrac{1}{2} = 1$$

$$\therefore \quad \text{Var}(X_k) = E(X_k^2) - E(X_k))^2 = 1 - 0 = 1$$

$$B_n = \text{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \text{Var}(X_i) \quad ;$$

[$\because$ $X_i$'s $(i = 1, 2, \ldots, n)$ are independent]

$$= \sum_{i=1}^{n} (1) = n$$

$$\therefore \quad \lim_{n\to\infty} \frac{B_n}{n^2} = \lim_{n\to\infty} \frac{1}{n} \to 0$$

Hence (Weak) Law of large numbers, holds for the sequence of independent r.v.'s $\{X_k\}$.

**Example 6·61.** *Let* $X_1, X_2, \ldots X_n$ *be jointly normal with* $E(X_i) = 0$ *and* $E(X_i^2) = 1$ *for all i and* $\text{Cov}(X_i, X_j) = \rho$ *if* $|j - i| = 1$ *and* $= 0$, *otherwise. Examine if WLLN holds for the sequence* $\{X_n\}$. ...(*)

**Solution.** We have:

$$\text{Var}(X_i) = (X_i^2) - [E(X_i)]^2 = 1, \quad (i = 1, 2, \ldots, n).$$

$$E(S_n) = E\left(\sum_{i=1}^{n} X_i\right) = 0$$

$$\text{Var}(X_n) = \text{Var}\left(\sum_{i=1}^{n} X_i\right)$$

$$= \sum_{i=1}^{n} \text{Var}\, X_i + 2 \sum_{i<j=1}^{n} \text{Cov}(X_i, X_j)$$

$$= n + 2 \cdot (n-1)\,\rho$$ [On using (*)]

Since $X_1, X_2, \ldots, X_n$ are jointly normal,

$$S_n = \sum_{i=1}^{n} X_i \sim N(0, \sigma^2) \text{ where } \sigma^2 = n + 2(n-1)\,\rho > 0\,;$$

Taking $Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i = \frac{S_n}{n}$, we have

$$E\left[\frac{Y_n^2}{1+Y_n^2}\right] = E\left[\frac{S_n^2}{n^2+S_n^2}\right]$$

$$= \int_{-\infty}^{\infty} \left(\frac{x^2}{n^2+x^2}\right) \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x^2/2\sigma^2}\, dx$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \frac{\sigma^2 y^2}{n^2+\sigma^2 y^2} \cdot e^{-y^2/2}\, dy\ ; \qquad \left(\frac{x}{\sigma} = y\right)$$

$$= \frac{2}{\sqrt{2\pi}} \{n + 2(n-1)\rho\} \cdot \int_0^{\infty} \frac{y^2}{n^2 + y^2 \{n + 2(n-1)\rho\}} \cdot e^{-y^2/2}\, dy$$

$$\le \frac{n + 2(n-1)\rho}{n^2} \cdot \frac{2}{\sqrt{2\pi}} \int_0^{\infty} y^2 \cdot e^{-y^2/2}\, dy$$

$$\rightarrow 0 \text{ as } n \rightarrow \infty$$

$$\therefore \qquad \lim_{n\rightarrow\infty} E\left|\frac{Y_n^2}{1+Y_n^2}\right| \rightarrow 0$$

Hence by Theorem 6·33, WLLN holds for $\{X_n\}$, *i.e.*, $\frac{S_n}{n} \xrightarrow{p} 0$ as $n \rightarrow \infty$.

**6·16. Borel-Cantelli Lemma.** (*Zero-One Law*). Let $A_1, A_2$ ... be a sequence of events. Let $A$ be the eent "that an infinite number of $A_n$ occur. That is $\omega \in A$ if $\omega \in A_n$ for an infinite number of values of $n$ (but not necessarily every $n$). But the set of such $\omega$ is precisely lim sup $A_n$, *i.e.*, $\overline{\lim}\, A_n$. Thus the event $A$, that an infinite number of $A_n$ occur is just $\overline{\lim}\, A_n$. Sometimes we are interested in the probability that an infinite number of the events $A_n$ occur. Often this question is answered by means of the Borel-Cantelli lemma or its converse.

**Theorem 6·34.** (*Borel-Cantelli Lemma*). *Let $A_1, A_2$, ... be a sequence of events on the probability space $(S, B, P)$ and let $A = \overline{\lim}\, A_n$.*

if $\sum_{n=1}^{\infty} P(A_n) < \infty$, *then* $P(A) = 0$.

In other words, this states that if $\Sigma\, P(A_n)$ converges then with probability one, only a finite number of $A_1, A_2, \ldots$ can occur.

**Proof.** Since $\quad A = \overline{\lim}\, A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$

we have $A \subset \bigcup_{m=n}^{\infty} A_m$, for every $n$.

Thus for each $n$,

$$P(A) \leq \sum_{m=n}^{\infty} P(A_m)$$

Since $\sum_{n=1}^{\infty} P(A_n)$ is convergent (given), $\sum_{m=n}^{\infty} P(A_m)$, being the remainder term of a convergent series, tends to zero as $n \to \infty$.

$\therefore$ $$P(A) \leq \sum_{m=n}^{\infty} P(A_m) \to 0 \text{ as } n \to \infty.$$

Thus $P(A) = 0$ as required.

The result just proved does not require events $A_1, A_2, \ldots$ considered to be independent. For the converse result it is necessary to make this further assumption.

**Theorem 6·35. Borel-Cantelli Lemma (Converse).** *Let $A_1, A_2, \ldots$ be independent events on $(S, \mathbf{B}, \mathbf{P})$, $A$ equal to $\overline{\lim}\ A_n$.*

*If $\sum_{n=1}^{\infty} P(A_n) = \infty$, then $P(A) = 1$.*

**Proof.** Writing, in usual notation, $\overline{A_n}$ for the complement $S - A_n$ of $A_n$, we have for any $m, n$ $(m > n)$.

$$\bigcap_{k=n}^{\infty} \overline{A_k} \subset \bigcap_{k=n}^{m} \overline{A_k}$$

$$P\left(\bigcap_{k=n}^{\infty} \overline{A_k}\right) \leq P\left(\bigcap_{k=n}^{m} \overline{A_k}\right)$$

$$= \prod_{k=n}^{m} P(\overline{A_k}),$$

because of the fact that if $(A_n, A_{n+1}, \ldots, A_m)$ are independent events, so are $(\overline{A_n}, \overline{A_{n+1}}, \ldots, \overline{A_m})$.

Hence $$P\left(\bigcap_{k=n}^{\infty} \overline{A_k}\right) \leq \prod_{k=n}^{m} [1 - P(A_k)]$$

$$\leq \prod_{k=n}^{m} e^{-P(A_k)}$$

$$[\because 1 - x \leq e^{-x} \text{ for } x \geq 0]$$

$$= e^{-\prod_{k=n}^{m} P(A_k)} \quad \forall\ m$$

Since $$\sum_{k=1}^{\infty} P(A_k) = \infty, \quad \sum_{k=n}^{m} P(A_k) \to \infty \text{ as } m \to \infty$$

and hence $e^{-\sum_{k=n}^{m} P(A_k)}\ \forall\ m \to 0$ as $m \to \infty$.

$\therefore$ $$P\left(\bigcap_{k=n}^{\infty} \overline{A_k}\right) = 0 \qquad \ldots(*)$$

But $$A = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

$\therefore$ $$\overline{A} = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \overline{A_k} \qquad \text{(De-Morgan's Law)}$$

$\Rightarrow$ $$P(\overline{A}) \leq \sum_{n=1}^{\infty} P\left(\bigcap_{k=n}^{\infty} \overline{A_k}\right) = 0 \qquad \text{[From (*)]}$$

Hence $$P(A) = 1 - P(\overline{A}) = 1, \text{ as required.}$$

If $A_1, A_2, \ldots$ are independent events it follows from Theorems 6·34 and 6·35 that the probability that an infinite number of them occur is either zero (when $\sum_{n=1}^{\infty} P(A_n) < \infty$) or one [when $\sum_{n=1}^{\infty} P(A_n) = \infty$]. This statement is a special case of so-called "*Zero one law*" which we now state.

**Theroem 6·36.** (*Zero One Law*): *If $A_1, A_2, \ldots$ are independent and if E belongs to the σ-field generated by the class $(A_n, A_{n+1}, \ldots)$ for every n, then P (E) is zero or one.*

**Example 6·62.** *What is the probability that in a sequence of Bernoulli trials with probability of success p for each trial, the pattern SFS appears infinitely often ?*

**Solution.** Let $A_k$ be the event that the trial number $k$, $K + 1$, $k + 2$ produce the sequence SFS ($k = 0, 1, 2, \ldots$). The events $A_k$ are not mutually independent but the sequence $A_1, A_4, A_7, A_{10}, \ldots$ contains only mutually independent events (since no two depend on the outcome of the same trials).

$$p_k = P(A_k) = P(SFS) = p^2 q, \qquad (q = 1 - p)$$

is independent of $k$, and hence the series

$$p_1 + p_4 + p_7 + \ldots, \text{ diverges}$$

Hence by B.C.T. (converse) the pattern SFS appears infinitely often with probability one.

**Example 6·63.** *A bag contains one black ball and white balls. A ball is drawn at random. If a white ball is drawn, it is returned to the bag together with an additional white ball. If the black ball is drawn, it alone is returned to the bag.*

*Let $A_n$ denote the event that the black ball is not drawn in the first n trials. Discuss the converse to Borel-Cantelli Lemma with reference to events $A_n$.*

**Solution.** $A_n$ = The event that blackball is not drawn in the first $n$ trials. ...(*)

= The event that each of the first $n$ trials resulted in the draw of a white ball.

$\Rightarrow$ $$P(A_n) = P(E_1 \cap E_2 \cap \ldots \cap E_n),$$

where $E_i$ is the event of drawing a white ball in the $i$th trial.

$$\therefore\ P(A_n) = P(E_1)\, P(E_2 \mid E_1)\, P(E_3 \mid E_1 \cap E_2) \ldots P(E_n \mid E_1 \cap E_2 \ldots \cap E_{n-1})$$

$$= \frac{m}{m+1} \times \frac{m+1}{m+2} \times \ldots \times \frac{m+n-1}{m+n} = \frac{m}{m+n}$$

(Since if first ball drawn is white it is returned together with an additional white ball, *i.e.*, for the seconod draw the box contains 1 $b$, $m + 1$ $W$ balls and

$$\therefore \qquad P(E_2 \mid E_1) = \frac{m+1}{m+2}, \text{ and so on.}$$

$$\sum_{n=1}^{\infty} PA(A_n) = \sum_{n=1}^{\infty} \frac{m}{m+n} = m \sum_{n=1}^{\infty} \frac{1}{m+n}$$

$$= m \left[ \frac{1}{m+1} + \frac{1}{m+2} + \frac{1}{m+3} + \ldots \right]$$

$$= m \left[ \sum_{n=1}^{\infty} \frac{1}{n} - \left( \sum_{n=1}^{m} \frac{1}{n} \right) \right] \qquad \ldots(**)$$

But $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent, $\left( \because \sum_{n=1}^{\infty} \frac{1}{n^p} \text{ is convergent, iff } p > 1 \right)$

and $\qquad \sum_{n=1}^{m} \frac{1}{n}$ is finite,

$\therefore \qquad$ *R.H.S.* of (**) is divergent.

Hence $\qquad \sum_{n=1}^{\infty} P(A_n) = \infty$

From the definition of $A_n$ in (*) it is obvious that $A_n \downarrow$.

$$\therefore \qquad A = \overline{\lim_n}\, A_n = \limsup_n A_n = \varphi$$

$$\therefore \qquad P(A) = P(\varphi) = 0$$

This result is *inconsistent* with converse of Borel-Cantelli Lemma, the reason being that *the events* $A_n$ $(n = 1, 2, \ldots)$ *considered here are not independent,*

$$\because \quad P(A_i \cap A_j) = P(A_i) = \frac{m}{m+i} \neq P(A_i)\, P(A_j),$$

since for $\qquad (j > i)\, A_j \subset A_i$ as $A_n \downarrow$

## EXERCISE 6 (d)

1. State and prove Chebychev's inequality.
2. (a) A random variable $X$ has a mean value of 5 and variance of 3.
   (*i*) What is the least value of Prob $[\,|X - 5| < 3]$ ?
   (*ii*) What value of $h$ guarantees that Prob $[\,|X - 5| < h] \geq 0{\cdot}99$ ?
   (*iii*) What is the least value of Prob $(\,|X - 5| < 7{\cdot}5)$ ?

(*b*) A random variable $X$ takes the values $-1$, 1, 3, 5 with associated probabilities 1/6, 1/6, 1/6 and 1/2. Find by direct computation $P(|X-3| \geq 1)$. Find an upper bound to this probability by applying Chebychev's inequality.

(*c*) If $X$ denote the sum of the numbers obtained when two dice are thrown, use Chebychev's inequality to obtain an upper bound for $P\{|X-7|>4\}$. Compare this with the actualy probability.

3. (*a*) An unbaised coin is tossed 100 times. Show that the probability that the number of heads will be between 30 and 70 is greater than 0·93 .

(*b*) Within what limits will the number of heads lie, with 95 p.c. probability, in 1000 tosses of a coin which is practically unbised?

(*c*) A symmetric die is thrown 720 times. Use Chebychev's inequality to find the lower bound for the probability of getting 100 to 140 sixes.

(*d*) Use Chebychev's inequality to determine how many times a fair coin must be tossed in order that the probability will be at least 0·95 that the ratio of the number of heads to the number of tosses will be between 0·45 and 0·55 .

**[Delhi Univ. B.Sc. (Stat. Hons.), 1988]**

4. (*a*) If you wish to estimate the proportion of engineers and scientists who have studied probability theory and you wish your estimate to be correct within 2% with probability 0·95 or more, how large a sample would you take (*i*) if you have no idea what the true proportion is, (*ii*) if you are confident that the true proportion is less than 0·2 ? **[Burdwan Univ. (B.Sc. (Hons.), 1992]**

**Hint.** (*i*) $\varepsilon = 2\%$ or 0·02

$$P[|f_n - p| < 0\cdot02] \geq 1 - \frac{1}{4n\,(0\cdot02)^2} = 0\cdot95$$

$$\therefore \qquad 0\cdot05 = \frac{1}{4n\,(0\cdot02)^2} \quad \Rightarrow \quad n = 12{,}500$$

(*ii*) $p < 0\cdot2,\ P[\,|f_n - p| \leq \varepsilon\,] \geq 1 - \dfrac{p(1-p)}{n\,\varepsilon^2}$

Now $p(1-p) < 0\cdot16$, therefore $1 - \dfrac{0\cdot16}{n\,\varepsilon^2} = 0\cdot95$

Hence $\quad n = 50 \times 50 \times 20 \times \cdot16 = 8000$

(*b*) Let the sample mean of a random variable $X$ be $\bar{X}$ s.d.s. Then if at lease 99 per cent of the values of $X$ fall within $K$ standard deviations from the mean, find $K$.

5. (*a*) If $X$ is a r.v. such that $E(X) = 3$ and $E(X^2) = 13$, use Chebychev's inequality to determine a lower bound for $P(-2 < X < 8)$.

**[Delhi Univ. B.Sc. (Maths Hons.), 1990]**

**Hint.** $\mu_x = 3,\ \sigma_x^2 = 4 \ \Rightarrow\ \sigma_x = 2$. Chebychev's inequality gives

$P[\,|X-3| < 2k\,] \geq 1 - 1/k^2 \ \Rightarrow\ P(3-2k < X < 3+2k) \geq 1 - 1/k^2$

Now taking $k = 2\cdot5$, we get $P(-2 < X < 8) \geq 21/25$.

(*b*) State and prove Chebychev's inequality. Use it to prove that in 2000 throws with a coin the probability that the number of heads lies between 900 and 1100 is at least 19/20. **[Delhi Univ. B.Sc. (Maths Hons.), 1989]**

**6.** (*a*) A random variable $X$ has the density function $e^{-x}$ for $x \geq 0$. Show that Chebychev's inequality gives $P[|X-1|>2]<1/4$ and show that the actual probability is $e^{-3}$.

(*b*) Let $X$ have the p.d.f.:

$$f(x) = \frac{1}{2\sqrt{3}}, -\sqrt{3} < x < \sqrt{3}$$
$$= 0, \text{ elsewhere}.$$

Find the actual probability $P\left[\,|X-\mu| \geq \frac{3}{2}\sigma\,\right]$ and compare it with the upper bound obtained by Chebychev's inequality.

**7.** If $X$ has the distribution with p.d.f.

$$f(x) = e^{-x}, \; 0 \leq x < \infty,$$

use Chebychev's inequality to obtain a lower bound to the probability of the inequality $-1 \leq X \leq 3$, and compare it with actual value.

**8.** Explain the concept of "convergence in probability".

If $X_1, X_2, \ldots, X_n$ by r.v.s. with means $\mu_1, \mu_2, \ldots, \mu_n$ and standard deviations $\sigma_1, \sigma_2, \ldots, \sigma_n$ and if $\sigma_n \to 0$ as $n \to \infty$, show that $\bar{X}_n - \bar{\mu}_n$ converges to zero stochastically.

Hence show that if $m$ is the number of successes in $n$ independent trials, the probability of success at the $i$th trial being $p_i$ then $m/n$ converges in probability to $(p_1 + p_2 + \ldots + p_n)/n$.

**9.** (*a*) If $X_n$ takes the values 1 and 0 with corresponding probabilities $p_n$ and $1-p_n$, examine whether the weak law of large numbers can be applied to the sequence $\{X_n\}$ where the variables $X_n$, $n = 1, 2, 3, \ldots$ are independent.

(*b*) $\{X_i\}$, $i = 1, 2, \ldots$ is a sequence of independent random variables with expected value of $X_i$ equal to $m_i$ and variance of $X_i$ equal to $\sigma_i^2$. If $\frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2$ tends to zero as $n$ tends to infinity, show that the weak law of large numbers holds good to the sequence. **[Bombay Univ. B.Sc. (Stat.), 1992]**

**10.** $\{X_k\}$, $k = 1, 2, \ldots$ is a sequence of independent random variables each taking the values $-1, 0, 1$. Given that

$$P(X_k = 1) = \frac{1}{k} = P(X_k = -1), \; P(X_k = 0) = 1 - \frac{2}{k}.$$

Examine if the law of large numbers holds for this sequence.

**11.** (*a*) Derive Chebychev's inequality and show how it leads to the weak law of large numbers. Mention some important particular cases wherein the weak law of large numbers holds good.

(*b*) State and prove the weak law of large numbers. Deduce as a corollary Bernoulli theorem and comment on its applications.

(*c*) Examine whether the weak law of large numbers holds good for the sequence $X_n$ of independent random variables where

$$P\left(X_n = \frac{1}{\sqrt{n}}\right) = \frac{2}{3},\ P\left(X_n = -\frac{1}{\sqrt{n}}\right) = \frac{1}{3}$$

(*d*) $\{X_n\}$ is a sequence of independent random variables such that

$$P\left(X_n = \frac{1}{\sqrt{n}}\right) = p_n,\ P\left(X_n = 1 + \frac{1}{\sqrt{n}}\right) = 1 - p_n$$

Examine whether the weak law of large numbers is applicable to the sequence $\{X_n\}$

(*e*) If $X$ is a random variable and $E(X^2) < \infty$, then prove that

$$P\{|X| \geq a\} \leq \frac{1}{a^2} E(X^2),\ \text{for all } a > 0.$$

Use Chebyshev's inequality to show that for $n > 36$, the probability that in $n$ throws of a fair die, the number of sixes lies between $\frac{1}{6}n - \sqrt{n}$ and $\frac{1}{6}n + \sqrt{n}$ is at least $31/36$. **[Calcutta Univ. B.Sc. (Maths Hons.), 1991]**

**12.** Let $\{X_n\}$ be a sequence of mutually independent random variables such that

$$X_n = \pm 1 \text{ with probability } \frac{1 - 2^{-n}}{2}$$

and $$X_n = \pm 2^{-n} \text{ with probability } 2^{-n-1}$$

Examine whether the weak law of large numbers can be applied to the sequence $\{X_n\}$.

**13.** Examine whether the law of large numbers holds for the sequence $\{X_k\}$ of independent random variables defined by $P(X_k = \pm k^{-1/2}) = \frac{1}{2}$.

**14.** (*a*) State Khinchin's theorem.

(*b*) Let $X_1, X_2, X_3, \ldots$ be a sequence of independent and identically distributed r.v.'s, each uniform on $[0, 1]$. For the geometric mean

$$G_n = (X_1 X_2 \ldots X_n)^{1/n}$$

show that $G_n \xrightarrow{p} c$ for some finite number $c$. Find $c$.

**Hint.** $X \sim U[0, 1]$, let $Y = -\log X$;

Then $F_Y(y) = 1 - e^{-y}$; $\Rightarrow$ $f_Y(y) = e^{-y}$; $y \geq 0$.

$\therefore$ $Y_i = -\log X_i$, $(i = 1, 2, \ldots, n)$ are i.i.d. r.v.'s. with $E(Y_i) = 1$.

By Khinchin's theorem

$$\sum_{i=1}^{n} Y_n/n = -\left(\sum_{i=1}^{n} \log X_i/n\right) = -\log G_n \xrightarrow{p} E\,Y_i = 1 \quad \Rightarrow \quad G_n \xrightarrow{p} e^{-1} = c.$$

(*c*) Let $X_1, X_2, \ldots$ be i.i.d. r.v.'s with common p.d.f.

$$f(x) = \frac{1+\delta}{x^{2+\delta}}, x \geq 1, \ \delta > 0$$
$$= 0, \ x < 1$$

Discuss if WLLN holds for the sequence $\{X_n\}$.

**Hint.** $E X_i = (1+\delta)/\delta < \infty$ (finite). Hence by Khinchin's theorem

$$S_n/n = \Sigma X_i/n \xrightarrow{p} (1+\delta)/\delta \text{ as } n \to \infty.$$

**15.** Let $\{X_n\}$ be any sequence of r.v.'s Write $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Prove that a necessary and sufficient condition for the sequence $\{X_n\}$ to satisfy the weak law of large numbers is that

$$E\left[\frac{Y_n^2}{1+Y_n^2}\right] \to \text{ as } n \to \infty.$$

**Hint.** See Remark to Theorem 6·33.

**16.** State and prove Weak Law of Large Numbers. Determine whether it holds for the following sequence of independent random variables:

$$P(X_n = +1) = (1 - 2^{-n})/2 = P(X_n = -1)$$

**[Delhi Univ. B.Sc. (Maths Hons.), 1989]**

**17.** Let $X_1, X_2, \ldots$ be i.i.d. standard Cauchy variates. Show that the WLLN does not hold for the sequence $\{X_n\}$.

**Hint.** Use Theorem 6·33.

$$E\left(\frac{Y_n^2}{1+Y_n^2}\right) = E\left(\frac{S_n^2}{n^2+S_n^2}\right) = E\left(\frac{(S_n/n)^2}{1+(S_n/n)^2}\right)$$

$$= \int_{-\infty}^{\infty} \frac{x^2}{1+x^2} \cdot \frac{1}{\pi} \frac{1}{1+x^2}\, dx$$

$\left[\because \frac{S_n}{n} = \frac{X_1+X_2+\ldots+X_n}{n}\right.$ is also a standard Cauchy variate. See Remark 4, § 8·9·1]

$$= \frac{2}{\pi}\int_0^{\infty} \sin^2\theta\, d\theta = \frac{1}{2} \qquad (x = \tan\theta)$$

$$\Rightarrow \lim_{n\to\infty} E\left[\frac{Y_n^2}{1+Y_n^2}\right] \not\to 0 \Rightarrow \text{WLLN does not hold for } \{X_n\}.$$

**18.** (*a*) Examine if the WLLN holds for the sequence $\{X_n\}$ of i.i.d. r.v.'s with

$$P\left[X_i = (-1)^{k-1} . k\right] = \frac{6}{\pi^2 k^2};\ k = 1, 2, 3, \ldots, i = 1, 2, 3, \ldots$$

**[Delhi Univ. B.Sc. (Maths Hons.), 1990]**

**Hint.** $E(X_i) = \sum_{k=1}^{\infty} (-1)^{k-1} \cdot k \cdot \frac{6}{\pi^2 k^2} = \frac{6}{\pi^2} \sum_{k=1}^{\infty} (-1)^{k-1} \cdot (1/k)$

$$= \frac{6}{\pi^2} \left[1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \dots\right]$$

$$= \frac{6}{\pi^2} \cdot \log_e 2$$

[$\because$ The series in bracket is convergent by Leibnitz test for alternating series].

Hence by Khinchin's theorem, WLLN holds for the sequence $\{X_i\}$ of i.i.d. r.v.'s.

(*b*) The r.v.'s $X_1, X_2, \dots, X_n$ have equal expectations and finite variation. Is the weak law of large numbers applicable to this sequence if all the co-variances $\sigma_{ij}$ are negative? **[Delhi Univ. B.Sc. (Maths Hons.), 1987]**

**Hint.** $\frac{B_n}{n^2} = \frac{\text{Var}(X_1 + X_2 + \dots + X_n)}{n^2} = \frac{1}{n^2}\left[\sum_{i=1}^{n} \sigma_i^2 + 2 \sum_{i<j=1}^{n} \sigma_{ij}\right]$

$$< \frac{1}{n^2}\left(\sum_{i=1}^{n} \sigma_i^2\right) \to 0 \text{ as } n \to \infty \qquad (\because \sigma_i^2 \text{ are finite})$$

Hence WLLN holds.

**19.** State and prove Borel Cantelli Lemma.

In a sequence of Bernoulli trials let $A_n$ be the event that a run of $n$ consecutive successes occurs between the $2^n$th and $2^{n+1}$th trails. Show that if $p \geq \frac{1}{2}$, there is probability one that infinitely many $A_n$ occur; if $p < \frac{1}{2}$, then with probability one only finitely many $A_n$ occur.

**20.** Let $X_1, X_2, \dots$ be independent r.v.'s and $S_n = \sum_{k=1}^{n} X_k$. If $\sum_{n=1}^{\infty} \sigma^2_{X_n}$ converges, prove that the series $\Sigma (X_n - E X_n)$ converges in probability.

If $\frac{1}{b_n^2} \sum_{k=1}^{n} \sigma^2_{X_k} \to 0$, then prove that

$$\frac{S_n - ES_n}{b_n} \xrightarrow{P} 0.$$

Deduce Chebychev's inequality.

## 6·17. Probability Generating Function

**Definition.** *If $a_0, a_1, a_2, \dots$ is a sequence of real numbers and if*

$$A(s) = a_0 + a_1 s + a_2 s^2 + \dots = \sum_{i=0}^{\infty} a_i s^i \qquad \dots(6{\cdot}84)$$

*converges in some interval $-s_0 < s < s_0$, when the sequence is infinite then the function $A(s)$ is known as the generating function of the sequence $\{a_i\}$.*

The variable s has no significance of its own and is introduced to identify $a_i$ as the co-efficient of $s^i$ in the expansion of $A(s)$. If the sequence $\{a_i\}$ is bounded, then the comparison with the geometric series shows that $A(s)$ converges at least for $|s| < 1$.

In the particular case when $a_i$ is the probability that an integral valued discrete variable $X$ takes the value $i$,

*i.e.,* $a_i = p_i = P(X = i)$ ; $i = 0, 1, 2, \ldots$ with $\Sigma p_i = 1$, then the probability generating function, abbreviated as p.g.f., of r.v. $X$ is defined as :

$$P(s) = E(s^X) = \sum_{x=0}^{\infty} s^x . p_x \qquad \ldots(6{\cdot}85)$$

**Remarks.** 1. Obviously we have $P(1) = \sum_x p_x = 1$.

Thus a function $P(s)$ defined in (6·85) is a p.g.f. iff $p_x \geq 0 \ \forall \ x$ and $\sum_x p_x = 1$

2. *Relation between p.g.f. and m.g.f.*

Taking $s = e^t$ in (6·85), we get

$$P(e^t) = E(e^{tX}) = M_X(t). \qquad \ldots(6{\cdot}86)$$

*i.e.,* from p.g.f. we can obtain m.g.f. on replacing $s$ by $e^t$.

3. *Bivariate probability generating function.* The joint p.g.f. of two random variables $X_1$ and $X_2$ is a function of two variables $s_1$ and $s_2$ defined by :

$$P_{X_1, X_2}(s_1, s_2) = E(s_1^{x_1} . s_2^{x_2}) = \sum_{x_1} \sum_{x_2} s_1^{x_1} s_2^{x_2} . p(x_1, x_2) \qquad \ldots(6{\cdot}87)$$

Marginal p.g.f.'s can be obtained from (6·87) as given below.

$$P_{X_1}(s_1) = E(s_1^{x_1}) = P_{X_1, X_2}(s_1, 1) \ ; \ P_{X_2}(s_2) = E(s_2^{x_2}) = P_{X_1, X_2}(1, s_2) \qquad \ldots(6{\cdot}88)$$

4. Two r.v.'s are *independent* if and only if :

$$P_{X_1, X_2}(s_1, s_2) = P_{X_1}(s_1) . P_{X_2}(s_2). \qquad \ldots(6{\cdot}89)$$

The above concepts can be generalised to $n$ random variables

**Theorem 6·37.** *If X is a random variable which assumes only integral values with probability distribution*

$$P(X = k) = p_k \ ; \ k = 0, 1, 2, \ldots \ \textit{and} \ P(X > k) = q_k, \ k \geq 0$$

*so that* $q_k = p_{k+1} + p_{k+2} + \ldots = 1 - \sum_{i=0}^{k} p_i$ *and two generating functions are*

$$P(s) = p_0 + p_1 s + p_2 s^2 + \ldots$$

$$Q(s) = q_0 + q_1 s + q_2 s^2 + \ldots$$

*then for* $-1 < s < 1$, $Q(s) = \dfrac{1 - P(s)}{1 - s}$ $\qquad \ldots(6{\cdot}90)$

**Proof.** We have

$$q_{k-1} - q_k = p_k\,,\quad k \ge 1$$

$$\sum_{k=1}^{\infty} q_{k-1} s^k - \sum_{k=1}^{\infty} q_k s^k = \sum_{k=1}^{\infty} p_k s^k$$

$\Rightarrow$ $$s\,Q(s) - Q(s) + q_0 = P(s) - p_0$$

$\Rightarrow$ $$Q(s) = \frac{(p_0 + q_0) - P(s)}{(1-s)}$$

But $$p_0 + q_0 = p_0 + p_1 + p_2 + \ldots = 1$$

Hence the theorem.

**Theorem 6·38.** *For a random variable X, which assumes only integral values, the expectation E (X) can be calculated either from the probability distribution* $P(X = i) = p_i$ *or in terms of*

$$q_k = p_{k+1} + p_{k+2} + \ldots$$

*Thus* $$E(X) = \sum_{i=1}^{\infty} i p_i = \sum_{k=0}^{\infty} q_k$$

*In terms of the gererating functions*

$$E(X) = P'(1) = Q(1) \qquad \ldots(6\cdot91)$$

**Proof.** $P(s) = \sum_{k=0}^{\infty} p_k s^k$. If $E(X)$ exists, then

$$E(X) = \sum_{k=1}^{\infty} k\,p_k$$

$$P'(s) = \sum_{k=1}^{\infty} k\,p_k\,s^{k-1} \Rightarrow P'(1) = \sum_{k=1}^{\infty} k\,p_k$$

$\therefore$ $$E(X) = P'(1)$$

We know that

$$Q(s)\,[1-s] = 1 - P(s)$$

Differentiating both sides w.r.t. $s$, we get

$$Q'(s)\,[1-s] - Q(s) = -P'(s) \qquad \ldots(*)$$

$\therefore$ $$Q(1) = P'(1)$$

Hence $$E(X) = P'(1) = Q(1)$$

**Theorem 6·39.** *If* $E(X^2) = \Sigma k^2 p_k$ *exists, then*

$$E(X^2) = P''(1) + P'(1) = 2Q'(1) + Q(1)$$

*and hence* $V(X) = 2Q'(1) + Q(1) - \{Q(1)\}^2 = P''(1) + P'(1) - \{P'(1)\}^2$ $\ldots(6\cdot92)$

**Proof.** $P(s) = \sum_{k=0}^{\infty} p_k s^k$, $P'(s) = \Sigma\, k\,p_k\,s^{k-1}$

$$sP'(s) = \Sigma k p_k s^k \text{. Differentiating again, we get}$$

$$P'(s) + sP''(s) = \Sigma k^2 p_k s^{k-1}$$

$$\therefore \quad P'(1) + P''(1) = \sum_{k} k^2 p_k = E(X^2) \qquad ...(**)$$

$$\therefore \quad Q''(s)[1-s] - 2Q'(s) = -P''(s) \qquad \text{[Differentiating (*) again]}$$

Putting $s = 1$, we get

$$2Q'(1) = P''(1). \text{ Substituting in (**), we get}$$

$$E(X^2) = P'(1) + P''(1) = Q(1) + 2Q'(1)$$

$$\therefore \quad \text{Var}(X) = E(X^2) - \{E(X)\}^2 = P''(1) + P'(1) - \{P'(1)\}^2$$

$$= 2Q'(1) + Q(1) - \{Q(1)\}^2$$

**6·17·1. Probability Generating Function for the sum of independent variables (Convolutions).** If $X$ and $Y$ are non-negative independent, integral valued discrete random variables with respective probability generating functions,

$$P(s) = \sum_{k=0}^{\infty} p_k s^k, \quad p_k = P(X = k)$$

$$R(s) = \sum_{k=0}^{\infty} r_k s^k, \quad r_k = P(Y = k),$$

it is possible to deduce the probability generating function for the variable $Z = X + Y$, which is also clearly integral valued, in terms of $P(s)$ and $Q(s)$.

Let $w_k$ denote $P(Z = k)$. The event $Z = k$ is the union of the following mutually exclusive events,

$$(X = 0 \cap Y = k), (X = 1 \cap Y = k-1), (X = 2 \cap Y = k-2), ..., (X = k \cap Y = 0)$$

and

since the variables $X$ and $Y$ are independent, each joint probability is the product of the appropriate individual probabilities. Therefore the distribution $w_k = P(Z = k)$ is given by

$$w_k = p_0 r_k + p_1 r_{k-1} + p_2 r_{k-2} + \ldots + p_k r_0 \text{ for all integral } k \geq 0$$

The new sequence of probabilities $\{w_k\}$ defined in terms of the sequences $\{p_k\}$ and $\{r_k\}$ is called the convolution of these sequences and is denoted by

$$\{w_k\} = \{p_k\} * \{r_k\} \qquad ...(6·93)$$

**Theorem 6·40.**

$$\mu'_{(r)} = E\left[X(X-1)\ldots(X-r+1)\right] = \left[\frac{\partial^r}{\partial s^r} P(s)\right]_{s=1}$$

**Proof.** Differentiating (6·85) partially $r$ times w.r.t. $s$, we get

$$\frac{\partial^r P(s)}{\partial s^r} = \sum_{x} x(x-1)(x-2)\ldots(x-r+1)\, s^{x-r} p_x$$

$$\left[\frac{\partial^r P(s)}{\partial s^r}\right]_{s=1} = \sum_x x(x-1)(x-2)\ldots(x-r+1)\,p_x = \mu'_{(r)}$$

**Theorem 6·41.** *If $\{p_k\}$ and $\{r_k\}$ are the sequences with the generating functions $P(s), R(s)$ and $\{w_k\}$ is their convolution, then $W(s) = P(s)\,R(s)$, where $W(s) = \Sigma\, w_k s^k$ is the generating function of the sum $X + Y$.*

**Proof.** Since the co-efficient of $s^k$ in the product $P(s)\,R(s)$ is

$$p_0 r_k + p_1 r_{k-1} + \ldots + p_{k-1} r_1 + p_k r_0 = w_k,$$

it follows that the probability generating function for $Z$, namely,

$W(S) = \sum\limits_{k=0}^{\infty} w_k s^k$ is equal to $P(s)\,R(s)$.

**Cor.** If $X_1, X_2, \ldots, X_n$ are independent integral–valued discrete variables with respective probability generating functions $P_1(s), P_2(s), \ldots, P_n(s)$ and if $Z = X_1 + X_2 + \ldots + X_n$, the probability generating function for $Z$ is given by

$$P_Z(s) = \prod_{i=1}^{n} P_i(s)$$

In particular, when $X_1, X_2, \ldots, X_n$ all have a common distribution and hence common probability generating function $P(s)$, we have

$$P_Z(s) = [P(s)]^n$$

**Example 6·64.** *Can $P(s) = 2/(1+s)$ be the p.g.f. of a r.v. $X$? Give reasons.*

**Solution.** We have $P(1) = 2/2 = 1$.

Also $\quad P(s) = \sum\limits_{r=0}^{\infty} p_r s^r = 2\,(1+s)^{-1}$

$$= 2\,(1 - s + s^2 - s^3 + \ldots)$$

$$= 2 \sum_{r=0}^{\infty} (-1)^r . s^r \qquad \ldots(*)$$

$\Rightarrow \qquad p_r = 2(-1)^r$

Hence $p^{2k+1} < 0$, *i.e.*, $p_1, p_3, p_5, \ldots$ are negative.

Since some co-efficient in $(*)$ are negative, $P(s)$ cannot be the p.g.f. of a r.v. $X$.

**Example 6·65.** *If $P(s)$ is the probability generating function for $X$, find the generating function for $(X-a)/b$.*

**Solution.** $\quad P(s) = E(s^X)$

P.G.F. for $\dfrac{X-a}{b} = E\left(s^{(x-a)/b}\right) = e^{-a/b} . E\left(s^{x/b}\right)$

$$= s^{-a/b} \cdot E\,[(s^{1/b})^x] = s^{-a/b}\,P\,(s^{1/b})$$

**Example 6·66.** *Let X be a random variable with generating function P (s). Find the generating function of (a) X + 1 (b) 2X.*

**Solution.** *(a)* $P(s) = \sum_{k=0}^{\infty} p_k s^k = E(s^X)$

$\therefore$ P.G.F. of $X + 1 = E(s^{X+1}) = s\,E(s^X) = s\,P(s)$

*(b)* P.G.F. of $2X = E(s^{2X}) = E[(s^2)^X] = P(s^2)$

**Example 6·67.** *Find the generating function of (a) $P(X \le n)$, (b) $P(X < n)$, and (c) $P(X = 2n)$.* **[Delhi Univ. M.Sc. (O.R.), 1989]**

**Solution.** *(a)* Let $X$ be an integral valued random variable with the probability distribution

$$P(X = n) = p_n \text{ and } P(X \le n) = q_n$$

so that $q_n = p_0 + p_1 + p_2 + \ldots + p_n$ ; $n = 0, 1, 2, \ldots$

$$\therefore \quad q_n - q_{n-1} = p_n\,,\ n \ge 1$$

$$\Rightarrow \quad \sum_{n=1}^{\infty} q_n s^n - \sum_{n=1}^{\infty} q_{n-1} s^n = \sum_{n=1}^{\infty} p_n s^n$$

$$\Rightarrow \quad Q(s) - q_0 - s\,Q(s) = P(s) - p_0$$

$$\Rightarrow \quad Q(s) = \frac{P(s) + q_0 - p_0}{1-s} = \frac{P(s)}{1-s} \qquad [\because q_0 = p_0]$$

*(b)* Let $q_n = P(X < n) = p_0 + p_1 + \ldots + p_{n-1}$, $q_n - q_{n-1} = p_{n-1}$, $n \ge 2$

$$\Rightarrow \quad \sum_{n=2}^{\infty} q_n s^n - \sum_{n=2}^{\infty} q_{n-1} s^n = \sum_{n=2}^{\infty} p_{n-1} s^n = s \sum_{n=1}^{\infty} p_n s^n$$

$$\Rightarrow \quad Q(s) - q_1 s - sQ(s) = sP(s) - sp_0 \qquad [\because q_0 = 0]$$

$$\Rightarrow \quad Q(s)\,[1 - s] = sP(s) - sp_0 + q_1 s \qquad [\because q_1 = p_0]$$

Hence $$Q(s) = \frac{sP(s)}{1-s}$$

*(c)* Let $P(X = 2n) = p_{2n}$

$$\therefore \quad Q(s) = \sum_{n=0}^{\infty} p_{2n} s^n = p_0 + p_2 s + p_4 s^2 + \ldots$$

$$2Q(s) = 2p_0 + 2p_2 s + 2p_4 s^2 + \ldots$$

$$= (p_0 + p_1 s^{1/2} + p_2 s + p_3 s^{3/2} + p_4 s^2 + \ldots) + (p_0 - p_1 s^{1/2} + p_2 s - p_3 s^{3/2} + \ldots)$$

$$= \sum_{k=0}^{\infty} p_k (s^{1/2})^k + \sum_{k=0}^{\infty} p_k (-s^{1/2})^k$$

$$= P(s^{1/2}) + P(-s^{1/2})$$

$$Q(s) = \frac{P(s^{1/2}) + P(s^{-1/2})}{2}$$

**Example 6·68.** *Let $\{X_k\}$ be mutually independent, each assuming the values $0, 1, 2, ..., a-1$ with probability $\frac{1}{a}$.*

*Let $S_n = X_1 + X_2 + \ldots + X_n$. Show that the generating function of $S_n$ is*

$$P(s) = \left[\frac{1-s^a}{a(1-s)}\right]^n$$

*and hence*

$$P(S_n = j) = \frac{1}{a^n}\sum_{\nu=0}^{\infty}(-1)^{\nu+j+a\nu}\binom{n}{\nu}\binom{-n}{j-a\nu}$$

**(Rajasthan Univ. M.Sc. 1992)**

**Solution.** As the $\{X_k\}$ are mutually independent variables and each $X_i$ assumes the same values $0, 1, 2, ..., a-1$, with the probability $\frac{1}{a}$, therefore each will have the same generating function and the generating function of $S_n$ will be the $n$th convolution of generating function of $X_1$. Now

$$P_{X_1}(s) = \sum_{k=0}^{a-1} p_k s^k = \frac{1}{a}[s^0 + s^1 + \ldots + s^{a-1}] = \frac{1-s^a}{a(1-s)}$$

$$\therefore \quad P_{S_n}(s) = \left[\frac{1-s^a}{a(1-s)}\right]^n$$

Now the probability $S_n = j$ is the co-efficient of $s^j$ in

$$\frac{1}{a^n}(1-s^a)^n(1-s)^{-n}$$

If we take $(\nu+1)$th term from $(1-s^a)^n$, then it will have the power of $s$ equivalent to $s^{a\nu}$ and hence to get the power of $s$ as $j$, we must take the term from $(1-s)^{-n}$ having the power of $s$ as $j - a\nu$.

$\therefore$ Required probability

$$= \frac{1}{a^n}\sum_{\nu=0}^{\infty}\binom{n}{\nu}(-1)^{\nu}\binom{-n}{j-a\nu}(-1)^{j-a\nu}$$

$$= \frac{1}{a^n}\sum_{\nu=0}^{\infty}(-1)^{j+\nu-a\nu}\binom{n}{\nu}\binom{-n}{j-a\nu}$$

$$= \frac{1}{a^n}\sum_{\nu=0}^{\infty}(-1)^{j+\nu+a\nu}\binom{n}{\nu}\binom{-n}{j-a\nu}$$

$$[\because (-1)^{2a\nu} = 1 \Rightarrow (-1)^{a\nu} = (-1)^{-a\nu}]$$

**Example 6·69.** *A random variable $X$ assumes the value $\lambda_1, \lambda_2, \ldots$ with probabilities $u_1, u_2, \ldots$, show that*

$$p_k = \frac{1}{k!}\sum_{j=0}^{\infty} u_j e^{-\lambda_j}(\lambda_j)^k\ ; \quad \lambda_j > 0,\ \Sigma u_j = 1$$

*is a probability distribution. Find its generating function and prove that its mean equals E (X) and variance equals V (X) + E (X).*

**Solution.**

$$\sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} \left[ \frac{1}{k!} \sum_{j=0}^{\infty} u_j e^{-\lambda_j} (\lambda_j)^k \right]$$

$$= \sum_{j=0}^{\infty} \left[ u_j e^{-\lambda_j} \sum_{k=0}^{\infty} (\lambda_j)^k / k! \right] = \sum_{j=0}^{\infty} u_j e^{-\lambda_j} e^{\lambda_j}$$

$$= \sum_{j=0}^{\infty} u_j = 1$$

Hence $p_k$ represents a prbability distribution.

Let $P(s)$ be the generating function of $p_k$, then

$$P(s) = \sum_{k=0}^{\infty} p_k s^k = \sum_{k=0}^{\infty} \left[ \frac{1}{k!} \sum_{j=0}^{\infty} u_j e^{-\lambda_j} (\lambda_j)^k s^k \right]$$

(Fubini's Theorem)

$$= \sum_{j=0}^{\infty} \left[ u_j e^{-\lambda_j} \sum_{k=0}^{\infty} (s\lambda_j)^k / k! \right]$$

$$= \sum_{j=0}^{\infty} u_j e^{-\lambda_j} e^{s\lambda_j} = \sum_{j=0}^{\infty} u_j e^{\lambda_j (s-1)}$$

Thus $$P(s) = \sum_{j=0}^{\infty} u_j \left[ 1 + \lambda_j (s-1) + \frac{\lambda_j^2 (s-1)^2}{2!} + \dots \right]$$

$$P'(1) = \sum_{j=0}^{\infty} u_j \lambda_j = E(X)$$

$$P''(s) = \sum_{j=0}^{\infty} u_j [\lambda_j^2 + \lambda_j^3 (s-1) + \dots]$$

$\therefore$ $$P''(1) = \sum_{j=0}^{\infty} \lambda_j^2 u_j = E(X^2)$$

$$V(p_k) = P''(1) + P'(1) - \{P'(1)\}^2 = E(X^2) + E(X) - \{E(X)\}^2$$
$$= E(X) + V(X)$$

## EXERCISE 6 (e)

1. *(a)* Define the probability generating function (p.g.f.) of a random variable.

*(b)* $X$ is a positive integral valued variable, such that $P(X = n) = p_n$, $n = 0, 1, 2, \dots$ Define the probability generating function $G(s)$ and the moment generating function $M(t)$ for $X$ and show that, $M(t) = G(e^t)$. Hence or otherwise prove that

$$E(X) = G'(1), \text{ var}(X) = G''(1) + G'(1) - [G'(1)]^2$$

*(c)* If $X$ and $Y$ are non-negative integral valued independent random variables with $P(s)$ and $Q(s)$ as their probability generating functions, show that their sum $X+Y$ has the p.g.f. $P(s)\,Q(s)$.

**2.** A test of the strength of a wire consists of bending and unbending until it breaks. Considering bending and unbending as two operations, let $X$ denote the random variable corresponding to the number of operations necessary to break the wire. If $P(X=r)=(1-p)\,p^{r-1}$ ; $r=1,2,3,\ldots$ and $0<p<1$, find the probability generating function of $X$.

**3.** Define the generating function $A(s)$ of the sequence $\{a_j\}$. Let $a_j$ be the number of ways in which the score $j$ can be obtained by throwing a die any number of times. Show that the g.f. of $\{a_j\}$ is $(1-s-s^2-s^3-s^4-s^5-s^6)^{-1}-1$.

**4.** Four tickets are drawn, one at a time with replacement, from a set of ten tickets numbered respectively 1, 2, 3, ..., 10, in such a way that at each draw each ticket is equally likely to be selected. What is the probability that the total of the numbers on the four drawn tickets is 20?

**Hint.** If $X_i$ denotes the number on the $i$th ticket then, for $i=1,2,3,4$, we observe that $X_i$ is an integral-valued variate with possible values 1, 2, 3, ..., 10, each having associated probability 1/10. Here each $X_i$ has

$$\text{g.f.}=\frac{1}{10}s+\frac{1}{10}s^2+\ldots+\frac{1}{10}s^{10}=\frac{1}{10}s\,(1-s^{10})\,(1-s)^{-1}$$

and, since the $X_i$'s are independent, it follows that the total of the numbers on the drawn tickets

$$Z=X_1+X_2+X_3+X_4$$

has probability generating function

$$\left\{\frac{1}{10}s\,(1-s^{10})\,(1-s)^{-1}\right\}^4=\frac{1}{10^4}s^4\,(1-s^{10})^4\,(1-s)^{-4}$$

The required probability is the co-efficient of $s^{16}$ in

$$\frac{1}{10^4}(1-s^{10})^4\,(1-s)^{-4}=\frac{63}{10{,}000}.$$

**5.** Find the generating functions of *(a)* $P(X\geq n)$, *(b)* $P(X>n+1)$.

**6.** *(a)* Obtain the generating function of $q_n$, the probability that in $n$ tosses of an ideal coin, no run of three heads occurs.

*(b)* Let $X$ be a non-negative integral-valued random variable with probability generating function $P(s)=\sum_{n=0}^{\infty}p_n s^n$. After observing $X$, conduct $X$ binomial trials with probability $p$ of success and let $Y$ denote the corresponding resulting number of successes.

Determine *(i)* the probability generating function of $Y$ and, *(ii)* probability generating function of $X$ given that $Y=X$.

7. In a sequence of Bernoulli trials, let $U_n$ be the probability that the first combination $SF$ occurs at trials number $(n-1)$ and $n$.

Find the generating function, mean and variance.

**Hint.** $U_2 = P(SF) = pq,\ U_3 = P(SSF) + P(FSF) = pq(p+q)$

$U_4 = (SSSF) + P(FSSF) + P(FFSF) = pq(p^2 + pq + q^2)$

In general $U_n = pq \sum_{k=0}^{n-2} p^k q^{n-2-k}$

$$= pq^{n-1}\left[1 + \frac{p}{q} + \left(\frac{p}{q}\right)^2 + \ldots + \left(\frac{p}{q}\right)^{n-2}\right]$$

$$= pq \cdot \frac{q^{n-1} - p^{n-1}}{q-p},$$

8. *(a)* In a sequence of Bernoulli trials, let $U_n$ be the probability of an even number of successes. Prove the recursion formula

$$U_n = qU_{n-1} + (1 - U_{n-1})\,p$$

From this derive the generating function and hence the explicit formula for $U_n$.

*(b)* A series of independent Bernoulli trials is performed until an uninterrupted run of $r$ successes is obtained for the first time where $r$ is a given positive integer. Assuming that the probability of a success in any trial is $p = 1 - q$, show that the probability generating function of the number of trials is

$$F(s) = \frac{p^r s^r (1 - ps)}{1 - s + q p^r s^{r+1}}$$

## ADDITIONAL EXERCISES ON CHPATER VI

1. *(a)* A horizontal line of length '5' units is divided into two parts. If the first part is of length $X$, find $E(X)$ and $E[X(5-X)]$.

*(b)* Show that

$$E(X-\mu)^3 = E(X^3) - 3\mu\sigma^2 - \mu^3$$

where $\mu$ and $\sigma^2$ are the mean and variance of $X$ respectively.

2. *(a)* Two players $A$ and $B$ alternately roll a pair of fair dice. $A$ wins if he gets six points before $B$ gets seven points and $B$ wins if he gets seven points before $A$ gets six points. If $A$ takes the first turn, find the probability that $B$ wins and the expected number of trials for $A$ to win.

*(b)* A box contains $2^n$ tickets among which ${}^nC_i$ tickets bear the numbers $i\ (i = 0, 1, 2, \ldots, n)$. A group of $m$ tickets is drawn. Let $S$ denote the sum of their numbers. Find the expectation and variance of $S$.

**Ans.** $\frac{1}{2}mn,\ \frac{1}{4}mn - \{mn(m-1)/4(2^n-1)\}$

3. In an objective type examination, consisting of 50 questions, for each question there are four answers of which only one is correct. A candidate scores 1 if he picks up the correct answer and $-1/3$ otherwise. If a candidate makes only a random choice in respect of each of the 50 questions, find his expected score and the variance of his score.

4. *(a)* A florist, in order to satisfy the needs of a number of regular and sophisticated customers, stocks a highly perishable flower. A dozen flowers cost Rs. 3 and sell for Rs. 10. Any flowers not sold the day they are stocked are worthless. Demand in dozens of flowers is as follows :

| Demand | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Probability | 0·1 | 0·2 | 0·3 | 0·2 | 0·1 | 0·1 |

*(i)* How many flowers should the florist stock daily in order to maximise the expected value of his net profit ?

*(ii)* Assuming that failure to satisfy any one customer's request will result in future lost profits amounting to Rs. 5·10 (goodwill cost), in addition to the lost profit on the immediate sale, how many flowers should the florist stock ?

*(iii)* What is the smallest goodwill cost of stocking five dozen flowers ?

**Hint.** For $i = 0, 1, 2, 3, 4, 5$, let $X_i$ be the random variable giving the florist's net profit, when he decides to stock '$i$' dozen flowers. Determine the probability function for each and the mean of each and pick up that '$i$' for which it is maximum.

**Ans.** *(i)* 3 dozen, *(ii)* 4 dozen and *(iii)* Rs. 2

5. Consider a sequence of Bernoulli trials with a constant probability $p$ of success in a single trial. Let $X_k$ denote the number of failures following the $(k-1)$th and preceding the $k$th success, and let $S_r = \sum_{k=1}^{r} X_k$.

Derive the probability distribution of $X_k$. Hence derive the probability distribution of $S_r$. Find $E(S_r)$ and $\text{Var}(S_r)$.

6. In the simplest type of weather forecasting — "rain" or "no rain" in the next 24 hours — suppose the probability of raining is $p$ $(> \frac{1}{2})$, and that a forecaster scores a point if his forecast proves correct and zero otherwise. In making $n$ independent forecasts of this type, a forecaster, who has no genuine ability, predicts "rain" with probability $\lambda$ and "no rain" with probability $(1 - \lambda)$. Prove that the probability of the forecast being correct for any one day is

$$[1 - p + (2p - 1)\lambda]$$

Hence derive the expectation of the total score $(S_n)$ of the forecaster for the $n$ days, and show that this attains its maximum value for $\lambda = 1$. Also, prove that

$$\text{Var}(S_n) = n[p - (2p - 1)\lambda][1 - p + (2p - 1)\lambda]$$

and thereby deduce that, for fixed $n$, this variance is maximum for $\lambda = \frac{1}{2}$.

**Hint.** $P(X_i = 1) = 1 - P(X_i = 0) = p\lambda + q(1-\lambda),\ S_n = \sum_{i=1}^{n} X_i,$

and the $X_i$ s being independent, the stated results follow.

**7.** In the simplest type of weather forecasting — rain or no rain in the next 24 hours — suppose the probability of raining is $p\ (> \frac{1}{2})$, and that a forecaster scores a point if his forecast proves correct and zero otherwise. In making $n$ indiependent forecasts of this type, a forecaster who has no genuine ability decides to allocate at random $r$ days to a "rain" forecast and the rest to "no rain". Find the expectation of his total score $\{S_n\}$ for the $n$ days and show that this attains its maximum value for $r = n$. What is the variance of $S_n$ ?

**Hint.** Let $X_i$ be a random variable such that

$$X_i = 1 \text{ if forecast is correct for ith day}$$
$$= 0 \text{ if forecast is incorrect for ith day, } (i = 1, 2, ..., n)$$

Then $P(X_i = 1) = \frac{r}{n},\ P(X_i = 0) = 1 - \frac{r}{n} = \frac{n-r}{n}$

$$\therefore \quad E(X_i) = p\left(\frac{r}{n}\right) + q\left(\frac{n-r}{n}\right) \text{ and } S_n = \sum_{i=1}^{n} X_i$$

But the $X_i$ 's are correlated random variables, so that for $i \neq j$,

$$E(X_iX_j) = P(X_i = 1 \cap X_j = 1)$$
$$= p \cdot \frac{r}{n}\left[p\left(\frac{r-1}{n-1}\right) + q\left(\frac{n-r}{n-1}\right)\right]$$
$$+ q\left(\frac{n-r}{n}\right)\left[p\left(\frac{r}{n-1}\right) + q\left(\frac{n-r-1}{n-1}\right)\right]$$

Hence $E(S_n) = np - (n-r)(p-q) < np$ for $p > q$ and $V(S_n) = npq$.

**8.** Let $n_1$ letters '$A$' and $n_2$ letters '$B$' be arranged at random in a sequence. A run is a succession of like letters preceded and followed by none or an unlike letter. Let $W$ be the total number of runs of '$A$' s and '$B$' s. Obtain expressions for Prob $\{W = r\}$, where $r$ is a given positive even integer and also when $r$ is odd.

Compute the expectation of $W$.

**9.** An urn contains $K$ varieties of objects in equal numbers. The objects are drawn one at a time and replaced before the next drawing. Show that the probability that $n$ and no less drawings will be required to produce objects of all varieties is

$$\sum_{r=0}^{k-1} (-1)^r\ {}^{k-1}C_r \left(\frac{k-1-r}{k}\right)^{n-1}$$

Hence or otherwise, find the expected number of drawings in a simple form.

10. An urn contains $a$ white and $b$ black balls. After a ball is drawn, it is to be returned to the urn if it is white, but if it is black, it is to be replaced by a white ball from another urn. Show that the probability of drawing a white ball after the foregoing operation has been repeated $x$ times is

$$1-\frac{b}{a+b}\left(1-\frac{1}{a+b}\right)^x$$

11. A box contains $k$ varieties of objects, the number of objects of each variety being the same. These objects are drawn one at a time and put back before the next drawing. Denoting by $n$ the smallest number of drawings which produce objects of all varieties, find $E(n)$ and $V(n)$.

12. There is a lot of $N$ objects from which objects are taken at random one by one with replacement. Prove that the expected value and variance of the least number of drawings needed to get $n$ different objects are respectively given by

$$N\left[\frac{1}{N}+\frac{1}{N-1}+\ldots+\frac{1}{N-n+1}\right]$$

and $$N\left[\frac{1}{(N-1)^2}+\frac{2}{(N-2)^2}+\ldots+\frac{n-1}{(N-n+1)^2}\right]$$

13. A large population consits of equal number of individuals of $c$ different types. Individuals are drawn at random one by one until at leasts one individual of each type has been found, whereupon sampling ceases. Show that the mean number of individuals in the sample is

$$c\left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{c}\right)$$

and the variance of the number is

$$c^2\left(1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots+\frac{1}{c^2}\right)-c\left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{c}\right)$$

14. (*a*) A point $P$ is taken at random in a line $AB$ of length $2a$, all positions of the point being equally likely. Show that the expected value of the area of the rectangle $AP.AB$ is $2a^2/3$ and the probability of the area exceeding $a^2/2$ $1/\sqrt{2}$.

(*b*) A point is chosen at random on a circle of radius $a$. Show that the expectation of its distance from another fixed point also on the circle is $4a/\pi$.

(*c*) Two points $P$ and $Q$ are selected at random in a square of side $a$. Prove that

$$E(|PQ|^2)=a^2/3$$

15. If the roots $x_1, x_2$ of the equation $x^2-ax+b=0$ are real and $b$ is positive but otherwise unknown, prove that

$$E(x_i)=\tfrac{1}{6}a \text{ and } E(x_2)=\tfrac{5}{6}a$$

16. ***(Banach's Match-box Problem).*** A certain mathematician always carries two match boxes (initially containing $N$ match-sides). Each time he wants a match-stick, he selects a box at random, inevitably a moment comes when he finds

a box empty. Show that the probability that there are exactly $r$ match-sticks in one box when the other box becomes empty is

$$^{2N-r}C_N \times \frac{1}{2^{2N-r}}$$

Prove also that the expected number of matches is

$$^{2N}C_N \times \frac{2N+1}{2^{2N}} - 1$$

**17.** $n$ couples procrete independetnly with no limits on family size. Births are single and independent and for the $i$th couple, the probability of a baby is $p_i$. The sex ratio $S$ is defined as

$$S = \frac{\text{Mean number of all boys}}{\text{Mean number of all children}}$$

Show that if all couples,

(*i*) Stop procreating on the birth of a boy, then

$$S = n \Big/ \sum_{i=1}^{n} \frac{1}{p^i}$$

(*ii*) Stop procreating on birth of a girl, then

$$S = 1 - \left[ n \Big/ \sum_{i=1}^{n} \frac{1}{q^i} \right], \text{ where } q_i = 1 - p_1$$

(*iii*) Stop procreating when they have children of both sexes, then

$$S = \left[ \sum_{i=1}^{n} \frac{1}{q_i} - \sum_{i=1}^{n} p_i \right] \Big/ \left[ \sum_{i=1}^{n} \frac{1}{p_i q_i} - n \right]$$

**18.** Show that if $X$ is a random variable such that $P(a \le X \le b) = 1$, then $E(X)$ and $\text{Var}(X)$ exist, and $a \le E(X) \le b$ and $\text{Var}(X) \le (b-a)^2/4$.

**19.** $(X, Y)$ is a two-dimensional discrete random variable with the possible values 0 and 1 for $X$, and also 0 and 1 for $Y$, and with the joint probabilities given by

| $Y$ \ $X$ | 0 | 1 |
|---|---|---|
| 0 | $p_{00}$ | $p_{10}$ |
| 1 | $p_{01}$ | $p_{11}$ |

Find the characteristic functions $\psi_1(t)$, $\psi_2(t)$ and $\xi(t_1, t_2)$ for $X, Y$ and $(X, Y)$ respectively and show that $\psi(t_1, t_2) = \psi_1(t_1)\psi_2(t_2)$ when $p_{00} p_{11} = p_{01} p_{10}$.

**20.** For a given sequence $\{X_n\}$ of r.v.'s,

$$\varphi_n(t, X_n) = (\sin nt)/nt,$$

determine the distribution function of $X_n$. Hence show that even though the sequence of characteristic functions $\varphi_n(t)$ converges to a limit $\varphi(t)$, the sequence of distribution functions does not converge to a distribution function. What is the condition that is violated here? [Indian Civil Services, 1984]

21. Two continuous variates $X$ and $Y$ have a joint p.d.f. with a joint characteristic function $\varphi(t_1, t_2)$. If $g_X(x)$ is the marginal density, show that $\mu'_r(x)$, the $r$th simple moment for the conditional distribution of $Y$ given $X = x$, satisfies the equation

$$\mu'_r(x) \cdot g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\partial^r \phi(t_1, 0)}{\partial t_2^r} e^{-it_1 x} dt_1$$

[Indian Civil Services, 1987]

22. Prove that the real part of a characteristic function is again a characteristic function. Prove further that if $\psi_1(t) = a_1(t) + ib_1(t)$ and $\psi_2(t) = a_2(t) + ib_2(t)$ are characteristic functions, then $a_1(t) a_2(t) - b_1(t) b_2(t)$ is a characteristic function.

23. Show that for any distribution

$$\int_{-\infty}^{\infty} \left(1 - \frac{x^2}{t^2}\right) dF(x) \le \int_{-t}^{t} dF(x)$$

and hence deduce $P[\,|X - E(X)| > k\sigma\,] \le \frac{1}{k^2}$, where $k > 0$ and $\text{Var}(X) = \sigma^2$.

24. (*a*) Let $X$ be a random variable with moment generating function $M(t), -h < t < h$. Prove that

$$P(X \ge a) \le e^{-at} M(t),\ 0 < t < h$$

and that

$$P(X \le a) \le e^{-at} M(t), -h < t < 0.$$

(*b*) Let $f(x, y) = xe^{-x(y+1)};\ x > 0, y > 0$
$= 0$, elsewhere

Find moment generating function of $Z = XY$.

**Hint.**
$$M_{XY}(t) = \int_0^{\infty}\int_0^{\infty} e^{txy} f(x, y)\, dx\, dy$$
$$= \int_0^{\infty} \left[ x e^{-x} \left\{ \int_0^{\infty} e^{-(1-t)xy} dy \right\} \right] dx\ ;\ 1 - t > 0$$
$$= \int_0^{\infty} x e^{-x} \frac{1}{(1-t)x} dx = \frac{1}{1-t};\ t < 1.$$

25. The probability of obtaining a 6 with a biased die is $p$, where $(0 < p < 1)$. Three players $A$, $B$ and $C$ roll this die in order, $A$ starting. The first

one to throw a 6 wins. Find the probability of winning for $A, B$ and $C$.

If $X$ is a random variable which takes the value $r$ if the game finishes at the $r$th throw, determine the probability generating function of $X$ and hence, or otherwise, evluate $E(X)$ and $\text{Var}(X)$.

**Hint.** Probabilities for the wins of $A, B$ and $C$ are $p/(1-q^3)$, $pq(1-q^3)$ and $pq^2/(1-q^3)$ respectively.

$$P(X=r)=pq^{r-1}, \text{ for } r \geq 1.$$

The probability generating function of $X$ is $P(s)=ps/(1-qs)$,

whence $\qquad E(X)=1/p$ and $\text{Var}(X)=1/qp^2$.

**26.** Define convergence in probability. Let $X_1, X_2, \ldots$ be i.i.d. variates with $f(x) = e^{-(x-1)}$, $x \geq 1$. Show that $Y_n \to 1$ in probability where $Y_n = \text{Min}(X_k)$; $1 \leq k \geq n$. **(Indian Civil Services, 1982)**

**27.** Let $\{X_n, n=1, 2, \ldots\}$ be a sequence of standardised variates and Corr $(X_m, X_n) = \exp[-|m-n|\alpha]$, $\alpha > 0$ and $m \neq n$. Show that W.L.L.N. holds for this sequence. **(Indian Civil Services, 1988)**

**28.** From the probability generating function (p.g.f.) of two random variables $X$ and $Y$ given by

$$P(s,t)=\exp[-\lambda-\mu-b+\lambda s+\mu t+bst],$$

(*i*) obtain the marginal p.g.f.'s and identify them,

(*ii*) obtain the p.g.f. of $X+Y$ and $P(X+Y)=0$, and

(*iii*) interpret the case $b=0$.

CHAPTER SEVEN

# Theoretical Discrete Probability Distributions

**7·0. Introduction.** In the previous chapters we have discussed in detail the frequency distributions. In the present chapter we will discuss theoretical discrete distributions in which variables are distributed according to some definite probability law which can be expressed mathematically. The present study will also enable us to fit a mathematical model or a function of the form $y = p(x)$ to the observed data.

We have already defined distribution function, mathematical expectation, m.g.f., characteristic function and moments. This prepares us for a study of theoretical distributions. This chapter is devoted to the study of univariate (except for the multinomial) distributions like Binomial, Poisson, Negative binomial Geometric, Hypergeometric, Multinomial and Power-series distributions.

**7·1. Bernoulli Distribution.** A random variable $X$ which takes two values 0 and 1, with probabilities $q$ and $p$ respectively, *i.e.*, $P(X = 1) = p$, $P(X = 0) = q$, $q = 1 - p$ is called a *Bernoulli variate* and is said to have a Bernoulli distribution.

**Remark.** Sometimes, the two values are +1, −1 instead of 1 and 0.

**7·1·1. Moments of Bernoulli distribution.** The $r^{th}$ moment about origin is

$$\mu_r' = E(X^r) = 0^r \cdot q + 1^r \cdot p = p \; ; r = 1, 2, \ldots \qquad \ldots(7\cdot1)$$

$$\mu_1' = E(X) = p, \; \mu_2' = E(X^2) = p$$

$$\mu_2 = \text{Var}(X) = p - p^2 = pq$$

The m.g.f. of Bernoulli variate is given by :

$$M_X(t) = e^{0 \cdot t} \times P(X = 0) + e^{1 \cdot t} \cdot P(X = 1) = q + pe^t \qquad \ldots(7\cdot1a)$$

**Remark. Degenerate Random Variable.** Sometimes we may come across a variate $X$ which is degenerate at a point '$c$', say, so that : $P(X = c) = 1$ and $= 0$ otherwise, *i.e.*, the whole mass of the variable is concentrated at a single point '$c$'.

$$\text{Since } P(X = c) = 1, \text{Var}(X) = 0.$$

Thus a degenerate r.v. $X$ is characterised by Var $(X) = 0$.

M.g.f. of degenerate r.v. is given by

$$M_X(t) = E(e^{tX}) = e^{tc} P(X = c) = e^{ct} \qquad \ldots(7\cdot1b)$$

**7·2. Binomial Distribution.** Binomial distribution was discovered by James Bernoulli (1654-1705) in the year 1700 and was first published posthumously in 1713, eight years after his death). Let a random experiment be performed repeatedly and let the occurrence of an event in a trial be called a success and its non-occurrence a failure. Consider a set of $n$ independent Bernoullian trials ($n$

being finite), in which the probability '$p$' of success in any trial is constant for each trial. Then $q = 1 - p$, is the probability of failure in any trial.

The probability of $x$ successes and consequently $(n - x)$ failures in $n$ independent trials, in a specified order (say) $SSFSFFFS...FSF$ (where $S$ represents success and $F$ failure) is given by the compound probability theorem by the expression :

$$P\,(SSFSFFFS...FSF) = P(S)P(S)P(F)P(S)P(F)P(F)P(F)P(S) \times \ldots \times P(F)P(S)P(F)$$

$$= p \cdot p \cdot q \cdot p \cdot q \cdot q \cdot q \cdot p \ldots q \cdot p \cdot q$$

$$= \underbrace{p \cdot p \ldots p}_{x \text{ factors}} \quad \underbrace{q \cdot q \cdot q \ldots q}_{(n-x) \text{ factors}} = p^x q^{n-x}$$

But $x$ successes in $n$ trials can occur in $\binom{n}{x}$ ways and the probability for each of these ways is $p^x q^{n-x}$. Hence the probability of $x$ successes in $n$ trials in *any order* whatsoever is given by the addition theorem of probability by the expression:

$$\binom{n}{x} p^x q^{n-x}$$

The probability distribution of the number of successes, so obtained is called the *Binomial probability distribution*, for the obvious reason that the probabilities of 0, 1, 2, ..., $n$ successes, *viz.*,

$q^n, \binom{n}{1} q^{n-1} p, \binom{n}{2} q^{n-2} p^2, \ldots, p^n$, are the successive terms of the binomial expansion $(q + p)^n$.

**Definition.** *A random variable $X$ is said to follow binomial distribution if it assumes only non-negative values and its probability mass function is given by*

$$P(X = x) = p(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} \; ; x = 0, 1, 2, \ldots, n \; ; q = 1 - p \\ 0, \text{ otherwise} \end{cases} \qquad \ldots(7\cdot2)$$

The two independent constants $n$ and $p$ in the distribution are known as the *parameters* of the distribution. '$n$' is also, sometimes, known as the degree of the binomial distribution.

Binomial distribution is a discrete distribution as $X$ can take only the integral values, *viz.*, 0, 1, 2,..., $n$. Any variable which follows binomial distribution is known as *binomial variate*.

We shall use the notation $X \sim B(n, p)$ to denote that the random variable $X$ follows binomial distribution with parameters $n$ and $p$.

The probability $p(x)$ in (7·2) is also sometimes denoted by $b(x, n, p)$.

**Remarks 1.** This assignment of probabilities is permissible because

$$\sum_{x=0}^{n} p(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (q + p)^n = 1$$

2. Let us suppose that $n$ trials constitute an experiment. Then if this experiment is repeated $N$ times, the *frequency function* of the binomial distribution is given by

$$f(x) = Np(x) = N\binom{n}{x} p^x q^{n-x}\ ; x = 0, 1, 2, ..., n \qquad ...(7.3)$$

and the expected frequencies of 0, 1, 2, ..., $n$ successes are the successive terms of the binomial expansion, $N(q+p)^n$, $q + p = 1$.

3. Binomial distribution is important not only because of its wide applicability, but because it gives rise to many other probability distributions. Tables for $p(x)$ are available for various values of $n$ and $p$.

4. **Physical conditions for Binomial Distribution.** We get the binomial distribution under the following experimental conditions.

(i) Each trial results in two mutually disjoint outcomes, termed as success and failure.

(ii) The number of trials '$n$' is finite.

(iii) The trials are independent of each other.

(iv) The probability of success '$p$' is constant for each trial.

The problems relating to tossing of a coin or throwing of dice or drawing cards from a pack of cards with replacement lead to binomial probabability distribution.

**Example 7·1.** *Ten coins are thrown simultaneously. Find the probability of getting at least seven heads.*

**Solution.** $p$ = Probability of getting a head $= \frac{1}{2}$

$q$ = Probability of not getting a head $= \frac{1}{2}$

The probability of getting $x$ heads in a random throw of 10 coins is

$$p(x) = \binom{10}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x} = \binom{10}{x}\left(\frac{1}{2}\right)^{10}\ ; x = 0, , 2, ..., 10$$

$\therefore$ Probability of getting at least seven heads is given by

$$P(X \geq 7) = p(7) + p(8) + p(9) + p(10)$$

$$= \left(\frac{1}{2}\right)^{10}\left\{\binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}\right\}$$

$$= \frac{120 + 45 + 10 + 1}{1024} = \frac{176}{1024}.$$

**Example 7·2.** *A and B play a game in which their chances of winning are in the ratio 3 : 2. Find A's chance of winning at least three games out of the five games played.* **[Burdwan Univ. B.Sc. (Hons.), 1993]**

**Solution.** Let $p$ be the probability that '$A$' wins the game. Then we are given $p = 3/5 \Rightarrow q = 1 - p = 2/5$.

Hence, by binomial probability law, the probability that out of 5 games played, $A$ wins '$r$' games is given by :

$$P(X=r)=p(r)=\binom{5}{r}\cdot(3/5)^r(2/5)^{5-r};\ r=0,1,2,\ldots,5$$

The required probability that '*A*' wins at least three games is given by :

$$P(X\geq 3)=\sum_{r=3}^{5}\binom{5}{r}\frac{3^r\cdot 2^{5-r}}{5^5}$$

$$=\frac{3^3}{5^5}\left[\binom{5}{3}2^2+\binom{5}{4}.3\times 2+1.3^2\times 1\right]=\frac{27\times(40+30+9)}{3125}=0\cdot68$$

**Example 7·3.** *If m things are distributed among 'a' men and 'b' women, show that the probability that the number of things received by men is odd, is*

$$\frac{1}{2}\left[\frac{(b+a)^m-(b-a)^m}{(b+a)^m}\right]$$

**(Nagpur Univ B.Sc., 1989, '93)**

**Solution.** $p$ = Probability that a thing is received by man $=\dfrac{a}{a+b}$, then

$q=1-p=1-\dfrac{a}{a+b}=\dfrac{b}{a+b}$, is the probability that a thing is received by woman.

The probability that out of $m$ things exactly $x$ are received by men and the rest by women, is given by

$$p(x)={}^mC_x p^x q^{m-x};\ x=0,1,2,\ldots,m$$

The probability $P$ that the number of things received by men is odd is given by

$$P=p(1)+p(3)+p(5)+\ldots={}^mC_1\cdot q^{m-1}\cdot p+{}^mC_3\cdot q^{m-3}\cdot p^3+{}^mC_5\cdot q^{m-5}\cdot p^5+\ldots$$

Now

$$(q+p)^m=q^m+{}^mC_1\cdot q^{m-1}\cdot p+{}^mC_2\cdot q^{m-2}p^2+{}^mC_3\cdot q^{m-3}\cdot p^3+{}^mC_4\cdot q^{m-4}\cdot p^4+\ldots$$

and

$$(q-p)^m=q^m-{}^mC_1\cdot q^{m-1}\cdot p+{}^mC_2\cdot q^{m-2}\cdot p^2-{}^mC_3\cdot q^{m-3}\cdot p^3+{}^mC_4\cdot q^{m-4}\cdot p^4-\ldots$$

$$\therefore\quad (q+p)^m-(q-p)^m=2\left[{}^mC_1\cdot q^{m-1}\cdot p+{}^mC_3\cdot q^{m-3}\cdot p^3+\ldots\right]=2P$$

But $\quad q+p=1$ and $q-p=\dfrac{b-a}{b+a}$

$$\therefore\quad 1-\left(\frac{b-a}{b+a}\right)^m=2P\ \Rightarrow\ P=\frac{1}{2}\left[\frac{(b+a)^m-(b-a)^m}{(b+a)^m}\right]$$

**Example 7·4** *An irregular six faced die is thrown and the expectation that in 10 throws it will give five even numbers is twice the expectation that it will give four even numbers. How many times in 10,000 sets of 10 throws each, would you expect it to give no even number.* **(Gujarat Univ. B.Sc. 1988)**

**Solution.** Let $p$ be the probability of getting an even number in a throw of a die. Then the probability of getting $x$ even numbers in ten throws of a die is

$$P(X=x)=\binom{10}{x}p^x q^{10-x};\ x=0,1,2\ldots10$$

We are given that

$$P(X = 5) = 2\,P(X = 4)$$

*i.e.,*
$$\binom{10}{5} p^5 q^5 = 2 \binom{10}{4} p^4 q^6$$

$$\Rightarrow \quad \frac{10\,!\,p}{5\,!\,5\,!} = 2\,\frac{10\,!\,q}{4\,!\,6\,!}$$

$$\Rightarrow \quad \frac{p}{5} = \frac{2q}{6} = \frac{q}{3}$$

$\therefore\ 3p = 5q = 5(1-p) \quad \Rightarrow \quad 8p = 5 \quad \Rightarrow \quad p = 5/8$ and $q = 3/8$

$$\therefore \quad P(X = x) = \binom{10}{x}\left(\frac{5}{8}\right)^x \left(\frac{3}{8}\right)^{10-x}$$

Hence the required number of times that in 10,000 sets of 10 throws each, we get no even number

$$= 10{,}000 \times P(X = 0) = 10{,}000 \times \left(\frac{3}{8}\right)^{10} = 1 \text{ (approx.)}$$

**Example 7·5** *In a precision bombing attack there is a 50% chance that any one bomb will strike the target. Two direct hits are required to destroy the target completely. How many bombs must be dropped to give a 99% chance or better of completely destroying the target ?* **[Gauhati Univ. M.A., 1992]**

**Solution.** We have :

$p$ = Probability that the bomb strikes the target = 50% = $\frac{1}{2}$. Let $n$ be the number of bombs which should be dropped to ensure 99% chance or better of completely destroying the target. This implies that "probability that out of $n$ bombs, at least two strike the target, is greater than 0·99".

Let $X$ be a r.v. representing the number of bombs striking the target. Then $X \sim B(n, p = \frac{1}{2})$ with

$$p(x) = P(X = x) = \binom{n}{x}\left(\frac{1}{2}\right)^x \cdot \left(\frac{1}{2}\right)^{n-x} = \binom{n}{x}\left(\frac{1}{2}\right)^n ;\ x = 0, 1, \ldots, n$$

We should have :

$$P(X \geq 2) \geq 0{\cdot}99$$

$$\Rightarrow \quad [1 - P(X \leq 1)] \geq 0{\cdot}99$$

$$\Rightarrow \quad [1 - \{p(0) + p(1)\}] \geq 0{\cdot}99$$

$$\Rightarrow \quad 1 - \left\{\binom{n}{0} + \binom{n}{1}\right\}\left(\frac{1}{2}\right)^n \geq 0{\cdot}99$$

$$\Rightarrow \quad 0{\cdot}01 \geq \frac{1+n}{2^n} \quad \Rightarrow \quad 2^n \times (0{\cdot}01) \geq 1 + n$$

$$\Rightarrow \quad 2^n \geq 100 + 100\,n \qquad \ldots(*)$$

By trial method, we find that the inequality (*) is satisfied by $n = 11$. Hence the minimum number of bombs needed to destroy the target completely is 11.

**Example 7·6.** *A department in a works has 10 machines which may need adjustment from time to time during the day. Three of these machines are old, each having a probability of 1/11 of needing adjustment during the day, and 7 are new, having corresponding probabilities of 1/21.*

*Assuming that no machine needs adjustment twice on the same day, determine the probabilits that on a particular day :*

*(i) just 2 old and no new machines need adjustment.*

*(ii) If just 2 machines need adjustment, they are of the same type.*

**(Nagpur Univ. B.E., 1989)**

**Solution.** Let $p_1$ = Probability that an old machine needs adjustment $= 1/11$

$\therefore \quad q_1 = 1 - p_1 = 10/11$

and $\quad p_2$ = Probability that a new machine needs adjustment = 1/21

$\therefore \quad q_2 = 1 - p_2 = 20/21$

Then $P_1(r)$ = Probability that '$r$' old machines need adjustment

$$= {}^3C_r\, p_1^r\, q_1^{3-r} = {}^3C_r\, (10/11)^{3-r}\, (1/11)^r$$

and $P_2(r)$ = Probability that '$r$' new machine need adjustment

$$= {}^7C_r\, p_2^r\, q_2^{7-r} = {}^7C_r\, (1/21)^r\, (20/21)^{7-r}$$

*(i)* The probability that just two old machines and no new machine need adjustment is given (by the compound probability theorem) by the expression :

$$P_1(2) \cdot P_2(0) = {}^3C_2 (1/11)^2 \cdot (10/11) \cdot (20/21)^7 = 0{\cdot}016$$

*(ii)* Similarly the probability that just 2 new machines and no old machine need adjustment is

$$P_1(0) \cdot P_2(2) = (10/11)^3 \cdot {}^7C_2\, (1/21)^2 \cdot (20/21)^5 = 0{\cdot}028$$

The probability that "If just two machines need adjustment, they are of the same type" is the same as the probability that "either just 2 old and no new or just 2 new and no old machines need adjustment".

$\therefore$ Required probability = 0·016 + 0·028 = 0·044

**7·2·1 Moments.** The first four moments about origin of binomial distribution are obtained as follows :

$$\mu_1' = E(X) = \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} = np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} q^{n-x} \, *$$

$$= np(q + p)^{n-1} = np \qquad (\because q + p = 1)$$

Thus the mean of the binomial distribution is $np$.

---

$$* \quad \binom{n}{x} = \frac{n}{x} \cdot \binom{n-1}{x-1} = \frac{n}{x} \cdot \frac{n-1}{x-1} \cdot \binom{n-2}{x-2}$$

$$= \frac{n}{x} \cdot \frac{n-1}{x-1} \cdot \frac{n-2}{x-2} \binom{n-3}{x-3}, \text{ and so on.}$$

$$\mu_2' = E(X^2) = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \{x(x-1) + x\} \frac{n(n-1)}{x(x-1)} \cdot \binom{n-2}{x-2} p^x q^{n-x}$$

$$= n(n-1)p^2 \left[ \sum_{x=2}^{n} \binom{n-2}{x-2} p^{x-2} q^{n-x} \right] + np$$

$$= n(n-1)p^2 (q+p)^{n-2} + np = n(n-1)p^2 + np$$

$$\mu_3' = E(X^3) = \sum_{x=0}^{n} x^3 \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \{x(x-1)(x-2) + 3x(x-1) + x\} p^x q^{n-x}$$

$$= n(n-1)(n-2)p^3 \sum_{x=3}^{n} \binom{n-3}{x-3} p^{x-3} q^{n-x}$$

$$+ 3n(n-1)p^2 \sum_{x=2}^{n} \binom{n-2}{x-2} p^{x-2} q^{n-x} + np$$

$$= n(n-1)(n-2)p^3 (q+p)^{n-3} + 3n(n-1)p^2 (q+p)^{n-2} + np$$

$$= n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np$$

Similarly

$$x^4 = x(x-1)(x-2)(x-3) + 6x(x-1)(x-2) + 7x(x-1) + x$$

Let $x^4 = Ax(x-1)(x-2)(x-3) + Bx(x-1)(x-2) + Cx(x-1) + x$

By giving to $x$ the values 1, 2 and 3 respectively, we find the values of arbitrary constants $A$, $B$ and $C$. Therefore,

$$\mu_4' = E(X^4) = \sum_{x=0}^{n} x^4 \binom{n}{x} p^x p^{n-x}$$

$$= n(n-1)(n-2)(n-3)p^4 + 6n(n-1)(n-2)p^3 + 7n(n-1)p^2 + np$$

[On simplification]

*Central Moments of Binomial Distribution :*

$$\mu_2 = \mu_2' - \mu_1'^2 = n^2p^2 - np^2 + np - n^2p^2 = np(1-p) = npq$$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$$

$$= \{n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np\} - 3\{n(n-1)p^2 + np\} np + 2(np)^3$$

$$= np[-3np^2 + 3np + 2p^2 - 3p + 1 - 3npq]$$

$$= np[3np(1-p) + 2p^2 - 3p + 1 - 3npq]$$

$= np[2p^2 - 3p + 1] = np\ (2p^2 - 2p + q) = npq\,(1 - 2p)$
$= npq\,[q + p - 2p] = npq\,(q - p)$

$\mu_4 = \mu_4' - 4\mu_3'\,\mu_1' + 6\mu_2'\,\mu_1'^2 - 3\mu_1'^4 = npq\,[1 + 3\ (n - 2)\,pq]$

[On simplification]

Hence

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{n^2p^2q^2(q-p)^2}{n^3p^3q^3} = \frac{(q-p)^2}{npq} = \frac{(1-2p)^2}{npq} \quad ...(7\cdot4)$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{npq\,\{1 + 3\,(n-2)\,pq\}}{n^2p^2q^2} = \frac{1 + 3\,(n-2)\,pq}{npq} = 3 + \frac{1-6pq}{npq} \quad ...(7\cdot5)$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}} = \frac{1-2p}{\sqrt{npq}},\ \gamma_2 = \beta_2 - 3 = \frac{1-6pq}{npq} \quad ...(7\cdot5\ a)$$

**Example 7·7** *Comment on the following :*
*The mean of a binomial distribution is 3 and variance is 4'.*

**Solution.** If the given binomial distribution has parameters $n$ and $p$, then we are given

$$\text{Mean} = np = 3 \quad ...(*)$$

and

$$\text{Variance} = npq = 4 \quad ...(**)$$

Dividing (**) by (*), we get $q = 4/3$,
which is impossible, since probability cannot exceed unity. Hence the given statement is wrong.

**Example 7·8.** *The mean and variance of binomial distribution are 4 and $\frac{4}{3}$ respectively. Find $P(X \geq 1)$.* **(Sardar Patel Univ. B.Sc. 1993)**

**Solution.** Let $X \sim B\ (n, p)$. Then we are given

$$\text{Mean} = E(X) = np = 4 \quad ...(*)$$

and $\text{Var}(X) = npq = \frac{4}{3}$

Dividing, we get

$$q = \tfrac{1}{3} \quad \Rightarrow \quad p = \tfrac{2}{3}$$

Substituting in (*), we get

$$n = \frac{4}{p} = \frac{4 \times 3}{2} = 6.$$

$P(X \geq 1) = 1 - P\ (X = 0) = 1 - q^n = 1 - (1/3)^6 = 1 - (1/729)$
$= 1 - 0{\cdot}00137 = 0{\cdot}99863$

**Example 7·9** *If $X \sim B\,(n, p)$, show that :*

$$E\left(\frac{X}{n} - p\right)^2 = \frac{pq}{n}\,;\ Cov\left(\frac{X}{n}, \frac{n-X}{n}\right) = -\frac{pq}{n}$$

**(Delhi Univ. B.Sc., 1989)**

**Solution.** Since $X \sim B(n, p)$, $E(X) = np$ and $\text{Var}(X) = npq$

$$\therefore \quad E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = p\,; \quad \text{Var}\left(\frac{X}{n}\right) = \frac{1}{n^2} \cdot \text{Var}(X) = \frac{pq}{n}$$

(i) $$E\left(\frac{X}{n} - p\right)^2 = E\left[\frac{X}{n} - E\left(\frac{X}{n}\right)\right]^2 = \text{Var}\left(\frac{X}{n}\right) = \frac{pq}{n}$$

(ii) $$\text{Cov}\left(\frac{X}{n}, \frac{n-X}{n}\right) = E\left[\left\{\frac{X}{n} - E\left(\frac{X}{n}\right)\right\}\left\{\frac{n-X}{n} - E\left(\frac{n-X}{n}\right)\right\}\right]$$

$$= E\left[\left(\frac{X}{n} - p\right)\left\{\left(1 - \frac{X}{n}\right) - (1-p)\right\}\right]$$

$$= E\left[\left(\frac{X}{n} - p\right)\left\{-\left(\frac{X}{n} - p\right)\right\}\right]$$

$$= -E\left(\frac{X}{n} - p\right)^2 = -\text{Var}\left(\frac{X}{n}\right) = -\frac{pq}{n}$$

**7·2·2 Recurrence Relation for the moments of Binomial Distribution.**
*(Renovsky Formula)*

By def.,

$$\mu_r = E\{X - E(X)\}^r = \sum_{x=0}^{n} (x - np)^r \binom{n}{x} p^x q^{n-x}$$

Differentiating with respect to $p$, we get

$$\frac{d\mu_r}{dp} = \sum_{x=0}^{n} \binom{n}{x} \Big[-nr(x-np)^{r-1} p^x q^{n-x}$$
$$+ (x-np)^r \left\{x p^{x-1} q^{n-x} - (n-x) p^x q^{n-x-1}\right\}\Big]$$

$$= -nr \sum_{x=0}^{n} \binom{n}{x} (x-np)^{r-1} p^x q^{n-x}$$
$$+ \sum_{x=0}^{n} \binom{n}{x} (x-np)^r p^x q^{n-x} \left\{\frac{x}{p} - \frac{n-x}{q}\right\}$$

$$= -nr \sum_{x=0}^{n} (x-np)^{r-1} p(x) + \sum_{x=0}^{n} (x-np)^r p(x) \frac{(x-np)}{pq}$$

$$= -nr \sum_{x=0}^{n} (x-np)^{r-1} p(x) + \frac{1}{pq} \sum_{x=0}^{n} (x-np)^{r+1} p(x)$$

$$\therefore \quad \frac{d\mu_r}{dp} = -nr\,\mu_{r-1} + \frac{1}{pq}\,\mu_{r+1}$$

$$\Rightarrow \quad \mu_{r+1} = pq\left[nr\,\mu_{r-1} + \frac{d\mu_r}{dp}\right] \qquad \ldots(7\cdot 6)$$

Putting $r = 1, 2$ and $3$ successively in (7·6), we get

$$\mu_2 = pq\left[n\mu_o + \frac{d\mu_1}{dp}\right] = npq \qquad (\because \mu_0 = 1 \text{ and } \mu_1 = 0)$$

$$\mu_3 = pq\left[2n\,\mu_1 + \frac{d\mu_2}{dp}\right] = pq\cdot\frac{d(npq)}{dp} = npq\frac{d}{dp}\{p(1-p)\}$$

$$= npq\frac{d}{dp}(p-p^2) = npq\,(1-2p) = npq\,(q-p)$$

and

$$\mu_4 = pq\left[3n\mu_2 + \frac{d\mu_3}{dp}\right] = pq\left[3n\cdot npq + \frac{d}{dp}\{npq\,(q-p)\}\right]$$

$$= pq\left[3n^2pq + n\frac{d}{dp}\{p(1-p)(1-2p)\}\right]$$

$$= pq\left[3n^2pq + n\frac{d}{dp}(p-3p^2+2p^3)\right]$$

$$= pq\,[3n^2pq + n(1-6p+6p^2)] = pq\,[3n^2pq + n(1-6pq)]$$

$$= npq\,[3npq + 1 - 6pq] = npq\,[1 + 3pq(n-2)]$$

**Example 7·10** *Show that the rth moment $\mu_r'$ about the origin of the binomial distribution of degree n is given by :*

$$\mu_r' = \left(p\frac{\partial}{\partial p}\right)^r (q+p)^n \qquad ...(*)$$

**[Patna Univ. B.Sc. (Hons.), 1993]**

**Solution.** We shall prove this result by using the principle of mathematical induction. We have

$$(q+p)^n = \sum_{x=0}^{n}\binom{n}{x}p^x q^{n-x} \Rightarrow \frac{\partial}{\partial p}(q+p)^n = \sum_{x=0}^{n}\binom{n}{x}q^{n-x}xp^{x-1}$$

$$\therefore\ \frac{\partial}{\partial p}(q+p)^n = p\sum_{x=0}^{n}\binom{n}{x}q^{n-x}xp^{x-1} = \sum_{x=0}^{n}\binom{n}{x}p^x q^{n-x}x = \mu_1'$$

Thus the result (*) is true for $r = 1$.

Let us now assume that the result (*) is true for $r = k$, so that

$$\left(p\frac{\partial}{\partial p}\right)^k (q+p)^n = \mu_k' = \sum_{x=0}^{n}\binom{n}{x}p^x q^{n-x}x^k \qquad ...(**)$$

Differentiate (**) partially w.r. to $p$ and multiply both sides by $p$ to get :

$$p\left(\frac{\partial}{\partial p}\right)\left[\left(p\frac{\partial}{\partial p}\right)^k (q+p)^n\right] = \sum_{x=0}^{n}\binom{n}{x}p^x q^{n-x}x^{k+1} = E\left(X^{k+1}\right)$$

$$\Rightarrow \qquad \left(p\frac{\partial}{\partial p}\right)^{k+1}(q+p)^n = \mu_{k+1}'$$

Hence if the result (*) is true for $r = k$, it is also true for $r = k+1$. It is already shown to be true for $k = 1$. Hence by the principle of mathematical induction, (*) is true for all positive integral values of $r$.

**7·2·3. Factorial Moments of Binomial Distribution.** The $r$th factorial moment of the Binomial distribution is:

$$\mu_{(r)}' = E\left\{X^{(r)}\right\} = \sum_{x=0}^{n} x^{(r)} p(x) = \sum_{x=0}^{n} x^{(r)} \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$= n^{(r)} p^r \sum_{x=r}^{n} \frac{(n-r)!}{(x-r)!(n-x)!} p^{x-r} q^{n-x} = n^{(r)} p^r (q+p)^{n-r}$$

$$= n^{(r)} p^r \qquad \ldots(7\cdot 7)$$

$$\mu_{(1)}' = E\left\{X^{(1)}\right\} = np = \text{Mean}$$

$$\mu_{(2)}' = E\left\{X^{(2)}\right\} = n^{(2)} p^2 = n(n-1)p^2$$

$$\mu_{(3)}' = E\left\{X^{(3)}\right\} = n^{(3)} p^3 = n(n-1)(n-2)p^3$$

Now $\mu_{(2)} = \mu_{(2)}' - \mu_{(1)}'^2 + \mu_{(1)}' = n^2p^2 - np^2 - n^2p^2 + np = npq$

$$\mu_{(3)} = \mu_{(3)}' - 3\mu_{(2)}'\mu_{(1)}' + 2\mu_{(1)}'^3 - 2\mu_{(1)}'$$

$$= n(n-1)(n-2)p^3 - 3n(n-1)p^2\,np + 2n^3p^3 - 2np = -2npq(1+p)$$

[On simplification]

**7·2·4. Mean Deviation About Mean of Binomial Distribution.** The mean deviation $\eta$ about the mean $np$ of the binomial distribution is given by

$$\eta = \sum_{x=0}^{n} |x - np|\; p(x) = \sum_{x=0}^{n} |x-np| \binom{n}{x} p^x q^{n-x},$$

($x$ being an integer)

$$= \sum_{x=0}^{np} -(x-np)\binom{n}{x} p^x q^{n-x} + \sum_{x=np}^{n} (x-np)\binom{n}{x} p^x q^{n-x}$$

$$= 2\sum_{x=np}^{n} (x-np)\binom{n}{x} p^x q^{n-x} \; *$$

$$= 2\sum_{\mu}^{n} (x-np)\binom{n}{x} p^x q^{n-x},$$

where $\mu$ is the greatest integer contained in $np+1$.

$$= 2\sum_{\mu}^{n} \left[\left\{xq - (n-x)p\right\}\binom{n}{x} p^x q^{n-x}\right]$$

$$= 2\sum_{\mu}^{n} \left[\frac{n!}{(x-1)!(n-x)!} p^x q^{n-x+1} - \frac{n!}{x!(n-x-1)!} p^{x+1} q^{n-x}\right]$$

---

* $\because \quad \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} = np$

$\Rightarrow \quad \sum_{x=0}^{n} (x-np)\binom{n}{x} p^x q^{n-x} = 0$

$$= 2 \sum_{x=\mu}^{n} \left[ t_{x-1} - t_x \right], \text{ where } t_x = \frac{n!}{x!\,(n-x-1)!} p^{x+1} q^{n-x}$$

$$= 2 \left[ t_{\mu-1} - t_n \right] = 2\, t_{\mu-1}$$

This is obtained by summing over $x$ and using $t_n = o$

$$\therefore \eta = 2t_{\mu-1} = 2 \frac{n!}{(\mu-1)!\,(n-\mu)!} \cdot p^{\mu} q^{n-\mu+1}$$

$$= 2npq \binom{n-1}{\mu-1} p^{\mu-1} q^{n-\mu} \qquad \ldots(7\cdot8)$$

**7·2·5. Mode of the Binomial Distribution.** We have

$$\frac{p(x)}{p(x-1)} = \binom{n}{x} p^x q^{n-x} \Big/ \binom{n}{x-1} p^{x-1} q^{n-x+1}$$

$$= \frac{n!}{(n-x)!\,x!} p^x q^{n-x} \Big/ \frac{n!}{(x-1)!\,(n-x+1)!} p^{x-1} q^{n-x+1}$$

$$= \frac{(n-x+1)p}{xq} = \frac{xq + (n-x+1)p - xq}{xq}$$

$$= 1 + \frac{(n+1)p - x(p+q)}{xq} = 1 + \frac{(n+1)p - x}{xq} \qquad \ldots(7\cdot9)$$

Mode is the value of $x$ for which $p(x)$ is maximum.

We discuss the following two cases :

**Case 1.** When $(n+1)p$ is not an integer

Let $(n+1)p = m + f$, where $m$ is an integer and $f$ is fractional such that $0 < f < 1$. Substituting in (7·9), we get

$$\frac{p(x)}{p(x-1)} = 1 + \frac{(m+f) - x}{xq} \qquad \ldots(*)$$

From (*), it is obvious that

$$\frac{p(x)}{p(x-1)} > 1 \text{ for } x = 0, 1, 2, \ldots, m$$

and $$\frac{p(x)}{p(x-1)} < 1 \text{ for } x = m+1, m+2, \ldots, n$$

$\Rightarrow$ $$\frac{p(1)}{p(0)} > 1, \frac{p(2)}{p(1)} > 1, \ldots, \frac{p(m)}{p(m-1)} > 1,$$

and $$\frac{p(m+1)}{p(m)} < 1, \frac{p(m+2)}{p(m+1)} < 1, \ldots, \frac{p(n)}{p(n-1)} < 1,$$

$$\therefore p(0) < p(1) < p(2) < \ldots < p(m-1) < p(m) > p(m+1) > p(m+2) > p(m+3) \ldots > p(n),$$

Thus in this case there exists unique modal value for binomial distribution and it is $m$, the integral part of $(n+1)p$.

**Case II.** *When $(n+1)p$ is an integer.*

Let $(n+1)p = m$ *(an integer).*

Substituting in (7·9), we get

$$\frac{p(x)}{p(x-1)} = 1 + \frac{m-x}{xq} \qquad \ldots(**)$$

From (**) it is obvious that

$$\left.\frac{p(x)}{p(x-1)}\right\} \begin{array}{l} > 1 \text{ for } x = 1, 2, \ldots, m-1 \\ = 1 \text{ for } x = m \\ < 1 \text{ for } x = m+1, m+2, \ldots, n \end{array}$$

Now proceeding as in case 1, we have :

$p(0) < p(1) < \ldots < p(m-1) = p(m) > p(m+1) > p(m+2) > \ldots > p(n)$

Thus in this case the distribution is bimodal and the two modal values are $m$ and $m-1$.

**Example 7·11.** *Determine the binomial distribution for which the mean is 4 and variance 3 and find its mode.* **(Madurai Kamraj Univ B.Sc. 1993)**

**Solution,** Let $X \sim B(n, p)$, then we are given that

$$E(X) = np = 4 \qquad ..(*)$$

and

$$\text{Var}(X) = npq = 3 \qquad ..(**)$$

Dividing (**) by (*), we get

$$q = \tfrac{3}{4} \quad \Rightarrow \quad p = 1 - q = \tfrac{1}{4}$$

Hence from (*), $\quad n = \frac{4}{p} = 16$

Thus the given binomial distribution has parameters $n = 16$ and $p = 1/4$.

**Mode.** We have $(n+1)p = 4{\cdot}25$, which is not an integer. Hence the unique mode of the binomial distribution is 4, the integral part of $(n+1)p$.

**Example 7·12.** *Show that for $p = 0{\cdot}50$, the binomial distribution has a maximum probability at $X = \frac{1}{2}n$, if $n$ is even, and at $X = \frac{1}{2}(n-1)$ as well as $X = \frac{1}{2}(n+1)$, if $n$ is odd.* **(Mysore Univ., B. Sc. 1991)**

**Solution.** Here we have to find the mode of the binomial distribution.

(i) Let $n$ be even $= 2m$, (say), $m = 1, 2, \ldots$

$\therefore$ If $p = 0{\cdot}5$, then $(n+1)p = (2m+1) \times \left(\frac{1}{2}\right) = m + 0{\cdot}5$

Hence in this case, the distribution is unimodal, the unique mode being at $X = m = n/2$.

(ii) Let $n$ be odd $= (2m+1)$, say. Then

$(n+1)p = (2m+2) \times \frac{1}{2} = m + 1$ (Integer)

$$= \frac{n-1}{2} + 1 = \frac{n+1}{2}$$

Since $(n+1)p$ is an integer, the distribution is bimodal, the two modes being $\frac{1}{2}(n+1)$ and $\frac{1}{2}(n+1)-1 = \frac{1}{2}(n-1)$.

**7·2·6. Moment Generating Function of Binomial Distribution.** Let $X$ be a variable following binomial distribution, then

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x q^{n-x} = \sum_{x=0}^{n} (pe^t)^x q^{n-x}\binom{n}{x} = (q+pe^t)^n \qquad ...(7·10)$$

M.G.F. about Mean of Binomial Distribution :

$$E\left\{e^{t(X-np)}\right\} = E(e^{tx}e^{-tnp}) = e^{-tnp}\cdot E(e^{tX}) = e^{-tnp}\cdot M_X(t)$$

$$= e^{-tnp}(q+pe^t)^n = (qe^{-pt}+pe^{tq})^n \qquad ...(7·11)$$

$$= \left[q\left\{1-pt+\frac{p^2t^2}{2!}-\frac{p^3t^3}{3!}+\frac{p^4t^4}{4!}-...\right\}\right.$$

$$\left.+\,p\left\{1+tq+\frac{t^2q^2}{2!}+\frac{t^3q^3}{3!}-...\right\}\right]^n$$

$$= \left[1+\frac{t^2}{2!}pq+\frac{t^3}{3!}pq(q^2-p^2)+\frac{t^4}{4!}pq(q^3+p^3)+...\right]^n$$

$$= \left[1+\left\{\frac{t^2}{2!}\cdot pq+\frac{t^3}{3!}\cdot pq(q-p)+\frac{t^4}{4!}pq(1-3pq)+...\right\}\right]^n$$

$$= \left[1+\binom{n}{1}\left\{\frac{t^2}{2!}\cdot pq+\frac{t^3}{3!}pq(q-p)+\frac{t^4}{4!}pq(1-3pq)+...\right\}\right.$$

$$\left.+\binom{n}{2}\left\{\frac{t^2}{2!}pq+\frac{t^3}{3!}pq(q-p)+...\right\}^2+...\right]$$

Now $\mu_2 = \text{Coefficient of } \dfrac{t^2}{2!} = npq$

$$\mu_3 = \text{Coefficient of } \frac{t^3}{3!} = npq(q-p)$$

$$\mu_4 = \text{Coefficient of } \frac{t^4}{4!} = npq(1-3pq)+3n(n-1)p^2q^2$$

$$= npq(1-3pq)+3n^2p^2q^2-3np^2q^2$$

$$= 3n^2p^2q^2+npq(1-6pq)$$

**Example 7·13** *$X$ is binomially distributed with parameters $n$ and $p$. What is the distrbution of $Y = n - X$?* **[Delhi Univ. B.Sc. (Maths Hons.), 1990]**

**Solution.** $X \sim B(n,p)$, represents the number of successes in $n$ independent trials with constant probability $p$ of success for each trial.

$\therefore\ Y = n - X$, represents the number of failures in $n$ independent trial with constant probability '$q$' of failure for each trial. Hence $Y = n - X \sim B(n, q)$

**Aliter** Since $X \sim B(n, p)$; $M_X(t) = E(e^{tX}) = (q + pe^t)^n$

$$\therefore\ M_Y(t) = E(e^{tY}) = E(e^{t(n-X)})$$
$$= e^{nt} \cdot E(e^{-tX}) = e^{nt}\ M_X(-t)$$
$$= e^{nt} \cdot (q + pe^{-t})^n$$
$$= \left[e^t\left(q + pe^{-t}\right)\right]^n = \left(p + qe^t\right)^n$$

Hence by uniqueness theorem of m.g.f., $Y = n - X \sim B(n, q)$

**Example 7·14.** *The m.g.f. of a r.v. X is* $\left(\frac{2}{3} + \frac{1}{3}e^t\right)^9$. *Show that :*

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = \sum_{x=1}^{5} \binom{9}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{9-x}$$

**[Delhi Univ. B.Sc. (Maths Hons.), 1989]**

**Solution.** Since $M_X(t) = \left(\frac{2}{3} + \frac{1}{3}e^t\right)^9 = \left(q + pe^t\right)^n$,

by uniqueness theorem of m.g.f. $X \sim B\left(n = 9,\ p = \frac{1}{3}\right)$

Hence $E(X) = \mu_x = np = 3$; $\sigma_X^2 = npq = 9 \times \frac{1}{3} \times \frac{2}{3} = 2$

$\mu \pm 2\sigma = 3 \pm 2 \times \sqrt{2} = 3 \pm 2 \times 1{\cdot}4 = (0{\cdot}2,\ 5{\cdot}8)$

$$\therefore\ P(\mu - 2\sigma < X < \mu + 2\sigma) = P(0{\cdot}2 < X < 5{\cdot}8) = P(1 \le X \le 5)$$
$$= \sum_{x=1}^{5} p(x) = \sum_{x=1}^{5} {}^nC_x p^x q^{n-x}$$
$$= \sum_{x=1}^{5} {}^9C_x\ (1/3)^x (2/3)^{9-x}$$

**7·2·7. Additive Property of Binomial Distribution.** Let $X \sim B(n_1, p_1)$ and $Y \sim B(n_2, p_2)$ be independent random variables. Then

$$M_X(t) = (q_1 + p_1 e^t)^{n_1},\ M_Y(t) = (q_2 + p_2 e^t)^{n_2} \qquad \ldots(*)$$

What is the distribution of $X + Y$?

We have

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t) \quad [\because X \text{ and } Y \text{ are independent}]$$
$$= (q_1 + p_1 e^t)^{n_1} \cdot (q_2 + p_2 e^t)^{n_2} \qquad \ldots(**)$$

Since (**) cannot be expressed in the form $(q + p\,e^t)^n$, from uniqueness theorem of m.g.f.'s it follows that $X + Y$ is not a binomial variate. Hence, *in general the sum of two independent binomial variates is not a binomial variate.*

*In other words, binomial distribution does not possess the additive or reproductive property.*

However, if we take $p_1 = p_2 = p$ (say), then from (**), we get

$$M_{X+Y}(t) = (q + pe^t)^{n_1+n_2},$$

which is the m.g.f. of a binomial variate with parameters $(n_1 + n_2, p)$. Hence, by uniqueness theorem of m.g.f.'s $X + Y \sim B(n_1 + n_2, p)$. Thus the *binomial distribution possesses the additive or reproductive property if* $p_1 = p_2$.

**Generalisation.** *If $X_i$, $(i = 1, 2, ..., k)$ are independent binomial variates with parameters $(n_i, p)$, $(i = 1, 2, ..., k)$ then their sum* $\sum\limits_{i=1}^{k} X_i \sim B\left(\sum\limits_{i=1}^{k} n_i,\ p\right)$.

The proof is left as an exercise to the reader.

**Example 7·15.** *If the independent random variables $X, Y$ are binomially distributed, respectively with $n = 3$, $p = 1/3$, and $n = 5$, $p = 1/3$, write down the probability that $X + Y \geq 1$.*

**Solution.** We are given

$$X \sim B(3, \tfrac{1}{3}) \text{ and } Y \sim B(5, \tfrac{1}{3}).$$

Since $X$ and $Y$ are independent binomial random variables, with $p_1 = p_2 = \frac{1}{3}$, by the additive property of binomial distribution, we get

$$X + Y \sim B(3 + 5, \tfrac{1}{3}), \text{ i.e., } X + Y \sim B(8, \tfrac{1}{3})$$

$$\therefore \quad P(X + Y = r) = {}^8C_r\left(\tfrac{1}{3}\right)^r \left(\tfrac{2}{3}\right)^{8-r} \qquad ...(*)$$

$$\begin{aligned} \text{Hence } P(X + Y \geq 1) &= 1 - P(X + Y < 1) \\ &= 1 - P(X + Y = 0) \\ &= 1 - \left(\tfrac{2}{3}\right)^8 \end{aligned}$$

**7·2·8. Characteristic Function of Binomial Distribution.**

$$\varphi_X(t) = E(e^{itX}) = \sum_{x=0}^{n} e^{itx} p(x) = \sum_{x=0}^{n} e^{itx} \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} e^{itx} \binom{n}{x} (pe^{it})^x q^{n-x} = (q + pe^{it})^n \qquad ...(7·12)$$

**7·2·9. Cummulants of the Binomial Distribution.** Cumulant generating function is

$$K_X(t) = \log M_X(t) = \log (q + pe^t)^n = n \log (q + pe^t)$$

$$= n \log \left[q + p\left(1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots\right)\right]$$

$$= n \log \left[1 + p\left(t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots\right)\right]$$

$$= n\left[p\left(t+\frac{t^2}{2!}+\frac{t^3}{3!}+\frac{t^4}{4!}+\ldots\right)-\frac{p^2}{2}\left(t+\frac{t^2}{2!}+\frac{t^3}{3!}+\ldots\right)^2\right.$$

$$\left.+\frac{p^3}{3}\left(t+\frac{t^2}{2!}+\frac{t^3}{3!}+\ldots\right)^3-\frac{p^4}{4}\left(t+\frac{t^2}{2!}+\frac{t^3}{3!}+\ldots\right)^4+\ldots\right]$$

Mean $= \kappa_1 =$ Coefficient of $t$ in $K_X(t) = np$

$$\mu_2 = \kappa_2 = \text{Coefficient of } \frac{t^2}{2!} \text{ in } K_X(t) = n(p-p^2) = np(1-p) = npq$$

The coefficient of $t^3$ in $K_X(t)$

$$= n\left[\frac{p}{3!}-\frac{p^2}{2!}\cdot 2\cdot\frac{1}{2!}+\frac{p^3}{3}\right] = \frac{np}{3!}(1-3p+2p^2)$$

$$\therefore\ \kappa_3 = \text{Coefficient of } \frac{t^3}{3!} \text{ in } K_X(t) = np(1-3p+2p^2)$$

$$= np(1-p)(1-2p) = npq(1-p-p) = npq(q-p)$$

$$\therefore\qquad \mu_3 = \kappa_3 = npq(q-p)$$

The Coefficient of $t^4$ in $K_X(t)$

$$= n\left[\frac{p}{4!}-\frac{p^2}{2!}\left(\frac{2}{3!}+\frac{1}{4}\right)+\frac{p^3}{3}\cdot\frac{3}{2!}-\frac{p^4}{4}\right]$$

$$= \frac{np}{4!}[1-7p+12p^2-6p^3]$$

$$\therefore\ \kappa_4 = \text{Co efficient of } \frac{t^4}{4!} \text{ in } K_X(t) = np(1-p)(1-6p+6p^2)$$

$$= npq[1-6p(1-p)] = npq(1-6pq)$$

$$\therefore\ \mu_4 = \kappa_4 + 3\kappa_2^2 = npq(1-6pq)+3n^2p^2q^2$$

$$= npq(1-6pq+3npq) = npq[1+3pq(n-2)]$$

**7·2·10. Recurrence Relation for Cumulants of Binomial Distribution.** By def.,

$$\kappa_r = \left[\frac{d^r}{dt^r}\log M_X(t)\right]_{t=0} = n\left[\frac{d^r}{dt^r}\log(q+pe^t)\right]$$

$$\frac{d\kappa_r}{dp} = n\left[\frac{d^r}{dt^r}\cdot\frac{d}{dp}\log(q+pe^t)\right]_{t=0} = n\left[\frac{d^r}{dt^r}\cdot\frac{(-1+e^t)}{q+pe^t}\right]_{t=0}$$

$$\kappa_{r+1} = n\left[\frac{d^{r+1}}{dt^{r+1}}\log(q+pe^t)\right]_{t=0}$$

$$= n\left[\frac{d^r}{dt^r}\cdot\frac{d}{dt}\log(q+pe^t)\right]_{t=0} = n\left[\frac{d^r}{dt^r}\left(\frac{pe^t}{q+pe^t}\right)\right]_{t=0}$$

$$= n\left[\frac{d^r}{dt^r}\left(1-\frac{q}{q+pe^t}\right)\right]_{t=0} = -nq\left[\frac{d^r}{dt^r}\left(\frac{1}{q+pe^t}\right)\right]_{t=0}$$

Hence

$$\kappa_{r+1} - pq\frac{d\kappa_r}{dp} = -nq\left[\frac{d^r}{dt^r}\left(\frac{1}{q+pe^t}\right)\right]_{t=0} - npq\left[\frac{d^r}{dt^r}\left(\frac{e^t-1}{q+pe^t}\right)\right]_{t=0}$$

$$= -nq\left[\frac{d^r}{dt^r}\left\{\frac{1+pe^t-p}{q+pe^t}\right\}\right]_{t=0}$$

$$= -nq\left[\frac{d^r}{dt^r}\left\{\frac{q+pe^t}{q+pe^t}\right\}\right]_{t=0} = -nq\left[\frac{d^r}{dt^r}(1)\right]_{t=0} = 0$$

$$\therefore \qquad \kappa_{r+1} = pq\frac{d\kappa_r}{dp} \qquad \text{...(7·13)}$$

In particular,

$$\kappa_2 = pq\cdot\frac{d\kappa_1}{dp} = pq\cdot\frac{d}{dp}(np) = npq \qquad (\because \kappa_1 = \text{mean} = np)$$

$$\kappa_3 = pq\cdot\frac{d\kappa_2}{dp} = pq\cdot\frac{d(npq)}{dp} = npq(q-p)$$

$$\kappa_4 = pq\cdot\frac{d\kappa_3}{dp} = pq\cdot\frac{d}{dp}\{npq(q-p)\}$$

$$= npq\frac{d}{dp}\{p(1-p)(1-2p)\}$$

$$= npq\cdot\frac{d}{dp}(p-3p^2+2p^3) = npq(1-6p+6p^2)$$

$$= npq[1-6p(1-p)] = npq(1-6pq)$$

### 7·2·11. Probability Generating Function of Binomial Distribution

$$P(s) = \sum_{\kappa=0}^{n} P(X=k)s^\kappa = \sum_{\kappa=0}^{n}\binom{n}{k}(ps)^\kappa q^{n-k} = (ps+q)^n \qquad \text{...(7·13 a)}$$

The fact that this generating function is $n$th power of $(q+ps)$ shows that $p(x) = \{b(x; n, p)\}$ is the distribution of the sum $S_n = X_1 + X_2 + \ldots + X_n$ of $n$ random variables with the common generating function $(q+ps)$. Each variable $X_i$ assumes the value 0 with probability $q$ and 1 with probability $p$.

Thus $$\{b(k; n, p)\} = \{b(k; 1, p)\}^{n*} \qquad \text{...(7·13b)}$$

Let $X$ and $Y$ be two independent random variables having $b(k; m, p)$ and $b(k; n, p)$ as their distributions, then

$$P_X(s) = (q+ps)^m \text{ and } P_Y(s) = (q+ps)^n$$

$$\therefore \quad P_{X+Y}(s) = (q+ps)^m(q+ps)^n = (q+ps)^{m+n}$$

$$\therefore \{b(k; m, p)\} * \{b(k; n, p)\} = \{b(k; m+n, p)\} \qquad \text{...(7·13c)}$$

Also $\mu_{(1)}' = [\, n(q + ps)^{n-1} p \,]_{s=1} = np$

$\mu_{(2)}' = [\, n(n-1)(q + ps)^{n-2} p^2 \,]_{s=1} = n(n-1)p^2$ and so on.

$$\mu_{(r)}' = [n(n-1)\dots(n-r+1)(q+ps)^{n-r}p^r]_{s=1}$$
$$= n(n-1)\dots(n-r+1)p^r$$

**Example 7·16** *Show that*

$$E\left(\frac{1}{X+a}\right) = \int_0^1 t^{a-1}\, G(t)\, dt, \quad a > 0 \qquad \dots(*)$$

*where $G(t)$ is the probability generating function of $X$.*

*Find it when $X \sim B(n, p)$,* and $a = 1$

**[Delhi Univ. (Stat Hons.) Spl Course, 1988]**

**Solution.** R.H.S. $= \int_0^1 t^{a-1} \cdot G(t)\, dt = \int_0^1 t^{a-1} \left(Et^X\right) dt$

$$= \int_0^1 \left\{ t^{a-1} \left( \sum_x p_x t^x \right) \right\} dt = \sum_x \left[ p_x \int_0^1 t^{x+a-1}\, dt \right]$$

$$= \sum_x p_x \cdot \frac{1}{(x+a)} = E\left(\frac{1}{X+a}\right)$$

If $X \quad B(n, p)$, then $G(t) = \sum_{x=0}^{n} t^x p_x = (q + pt)^n \qquad \dots(**)$

Hence taking $a = 1$ in (*) and using (**), we get :

$$E\left[\frac{1}{(X+a)}\right] = \int_0^1 (q + pt)^n\, dt = \left| \frac{(q+pt)^{n+1}}{(n+1)p} \right|_0^1 = \frac{1 - q^{n+1}}{(n+1)p}$$

**7·2·12. Recurrence Relation for the Probabilities of Binomial Distribution.** *(Fitting of Binomial Distribution).*

We have

$$\frac{p(x+1)}{p(x)} = \frac{\binom{n}{x+1} p^{x+1} q^{n-x-1}}{\binom{n}{x} p^x q^{n-x}}$$

$$= \frac{n-x}{x+1} \cdot \frac{p}{q} \qquad \text{(On simplification)}$$

$$p(x+1) = \left\{ \frac{n-x}{x+1} \cdot \frac{p}{q} \right\} p(x), \qquad \dots(7\cdot 14)$$

which is the required recurrence formula.

This formula provides us a very convenient method of graduating the given data by a binomial distribution. The only probability we need to calculate is $p(0)$

which is given by $p(0) = q^n$, where $q$ is estimated from the given data by equating the mean $\bar{x}$ of the distribution to $np$, the mean of the binomial distribution. Thus $\hat{p} = \bar{x}/n$.

The remaining probabilities, *viz.*, $p(1), p(2), \ldots$ can now be easily obtained from (7·14) as explained below :

$$p(1) = [p(x+1)]_{x=0} = \left(\frac{n-x}{x+1} \cdot \frac{p}{q}\right)_{x=0} p(0)$$

$$p(2) = [p(x+1)]_{x=1} = \left(\frac{n-x}{x+1} \cdot \frac{p}{q}\right)_{x=1} p(1)$$

$$p(3) = [p(x+1)]_{x=2} = \left(\frac{n-x}{x+1} \cdot \frac{p}{q}\right)_{x=2} p(2)$$

and so on.

**Example 7·17.** *Seven coins are tossed and number of heads noted. The experiment is repeated 128 times and the following distribution is obtained:*

| *No. of heads* | *0* | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *Total* |
|---|---|---|---|---|---|---|---|---|---|
| *Frequencies* | *7* | *6* | *19* | *35* | *30* | *23* | *7* | *1* | *128* |

*Fit a Binomial distribution assuming*

*(i) The coin is unbaised,*

*(ii) The nature of the coin is not known.*

*(iii) Probability of a head for four coins is 0·5 and for the remaining three coins is 0·45.*

**Solution.** In fitting Binomial distribution, first of all the mean and variance of the data are equated to $np$ and $npq$ respectively. Then the expected frequencies are calculated from these values of $n$ and $p$. Here $n = 7$ and $N = 128$.

**Case I.** When the coin is unbaised

$$p = q = \tfrac{1}{2}, \ (p/q = 1)$$

Now
$$p(0) = q^n = (\tfrac{1}{2})^7 = (1/128)$$

$$f(0) = Nq^n = 128\,(\tfrac{1}{2})^7 = 1$$

Using the recurrence formula, the various probabilities, *viz.*, $p(1), p(2), \ldots$ can be easily calculated as shown below.

| $x$ | $\dfrac{n-x}{x+1}$ | $\dfrac{n-x}{x+1} \cdot \dfrac{p}{q}$ | *Expected frequency* $f(x) = Np(x)$ |
|---|---|---|---|
| 0 | 7 | 7 | $f(0) = Np(0) = 1$ |
| 1 | 3 | 3 | $f(1) = 1 \times 7 = 7$ |

| | | | |
|---|---|---|---|
| 2 | $\frac{5}{3}$ | $\frac{5}{3}$ | $f(2) = 7 \times 3 = 21$ |
| 3 | 1 | 1 | $f(3) = 21 \times \frac{5}{3} = 35$ |
| 4 | $\frac{3}{5}$ | $\frac{3}{5}$ | $f(4) = 35 \times 1 = 35$ |
| 5 | $\frac{1}{3}$ | $\frac{1}{3}$ | $f(5) = 35 \times \frac{3}{5} = 21$ |
| 6 | $\frac{1}{7}$ | $\frac{1}{7}$ | $f(6) = 21 \times \frac{1}{3} = 7$ |
| 7 | | | $f(7) = 7 \times \frac{1}{7} = 1$ |

**Case II.** When the nature of the coin is not known, then

$$np = \frac{1}{N} \sum_{i=1}^{n} f_i x_i = \frac{433}{128} = 3 \cdot 3828 \,;\; n = 7$$

$\therefore$
$$p = 0 \cdot 48326 \text{ and } q = 0 \cdot 51674, \; (p/q = 0 \cdot 93521)$$
$$f(0) = Nq^7 = 128\,(0 \cdot 5167)^7 = 1 \cdot 2593 \text{ (using logarithms)}$$

| x | $\frac{n-x}{x+1}$ | $\frac{n-x}{x+1} \cdot \frac{p}{q}$ | Expected frequency $f(x) = Np(x)$ |
|---|---|---|---|
| 0 | 7 | 6·54647 | $f(0) = Np(0) = 1 \cdot 2593 \simeq 1$ |
| 1 | 3 | 2·80563 | $f(1) = 1 \cdot 2593 \times 6 \cdot 54647 = 8 \cdot 2438 \simeq 8$ |
| 2 | $\frac{5}{3}$ | 1·55868 | $f(2) = 2 \cdot 80563 \times 8 \cdot 2438 = 23 \cdot 129 \simeq 23$ |
| 3 | 1 | 0·93521 | $f(3) = 1 \cdot 55868 \times 23 \cdot 129 = 36 \cdot 05 \simeq 36$ |
| 4 | $\frac{3}{5}$ | 0·56113 | $f(4) = 0 \cdot 93521 \times 36 \cdot 05 = 33 \cdot 715 \simeq 34$ |
| 5 | $\frac{1}{3}$ | 0·31174 | $f(5) = 0 \cdot 56113 \times 33 \cdot 715 = 18 \cdot 918 \simeq 19$ |
| 6 | $\frac{1}{7}$ | 0·13360 | $f(6) = 0 \cdot 31174 \times 18 \cdot 918 = 5 \cdot 897 \simeq 6$ |
| 7 | | | $f(7) = 0 \cdot 13360 \times 5 \cdot 897 = 0 \cdot 788 \simeq 1$ |

The probability generating functions (p.g.f.), say $P_X(s)$ for the 4 coins and $P_Y(S)$ for the remaining 3 coins are given by

$$P_X(s) = (0 \cdot 50 + 0 \cdot 50\, s)^4, \; P_Y(s) = (0 \cdot 55 + 0 \cdot 45 s)^3 \qquad \text{...[c.f. 7·13 (a)]}$$

Since all the throws are independent, the p.g.f. $P_{X+Y}(s)$ for the whole experiment is given by

$$P_{X+Y}(s) = P_X(s)\, P_Y(s) \qquad \ldots[c.f.\ 7\cdot 13\ (b)]$$
$$= (0\cdot 50 + 0\cdot 50\, s)^4\, (0\cdot 55 + 0\cdot 45\, s)^3$$
$$= (0\cdot 0625 + 0\cdot 25\, s + 0\cdot 375\, s^2 + 0\cdot 25\, s^3 + 0\cdot 0625\, s^4)$$
$$\times\ (0\cdot 166375 + 0\cdot 408375\, s + 0\cdot 334125\, s^2 + 0\cdot 091125\, s^3)$$

Now $f(x) = N \times$ coefficient of $t^x$ in $P_{X+Y}(t)$

$\therefore\ f(0) = 128 \times \cdot 0625 \times \cdot 16637 = 1\cdot 13310$

$f(1) = 128\,\{\cdot 25 + \cdot 166375 + \cdot 408375 \times \cdot 0625\} = 8\cdot 5910$

$f(2) = 128\,\{\cdot 28396\} = 36\cdot 3470 \qquad f(5) = 128\,\{\cdot 14602\} = 18\cdot 6934$

$f(3) = 128\,\{\cdot 184117\} = 23\cdot 5669 \qquad f(6) = 128\,\{\cdot 04366\} = 5\cdot 5889$

$f(4) = 128\,\{\cdot 260570\} = 33\cdot 3529 \qquad f(7) = 128\,\{\cdot 005695\} = \cdot 72896$

**Example 7·18.** *Let X and Y be independent binomial variates, each with parameters n and p. Find P (X − Y = k).* **(Calcutta Univ. B.Sc., 1993)**

**Solution.** Since each of the variables $X$ and $Y$ takes the values $0, 1, 2, \ldots, n$, $Z = X - Y$ takes on the values $-n, -(n-1), \ldots, -1, 0, 1, \ldots, n$

$$P(Z = k) = \sum_{r=0}^{n} P(X = k + r \cap Y = r)$$
$$= \sum_{r=0}^{n} P(X = k + r) \cdot P(Y = r) \quad (\because X \text{ and } Y \text{ are independent}).$$
$$= \sum_{r=0}^{n} \binom{n}{k+r} p^{k+r} \cdot q^{n-k-r} \binom{n}{r} p^r q^{n-r}$$
$$= \sum_{r=0}^{n} \binom{n}{k+r} \binom{n}{r} p^{2r+k}\, q^{2n-2r-k} \qquad \ldots(*)$$

where $k = -n, -(n-1), \ldots, -2, -1, 0, 1, 2, \ldots, n$; and $q = 1 - p$.

In particular, we have :

$$P(Z = 0) = \sum_{r=0}^{n} \binom{n}{r}^2 \cdot p^{2r}\, q^{2n-2r}$$
$$P(Z = -n) = \sum_{r=0}^{n} \binom{n}{-n+r} \binom{n}{r} p^{2r-n}\, q^{3n-2r} = p^n q^n,$$

because we get the result when $r = n$ and for other values of $r < n$, $\binom{n}{-n+r}$ is not defined and hence taken as 0.

**Example 7·19.** *Find the m.g.f. of standard binomial variate $(X - np)/\sqrt{npq}$ and obtain its limiting form as $n \to \infty$. Also interpret the result.*

**[Delhi Univ. B.Sc. (Stat. Hons.) 1990, 85]**

**Solution.** We know that if $X \sim B(n, p)$, then

$$M_X(t) = (q + p e^t)^n$$

The m.g.f. of standard binomial variate.

$$Z = \frac{X - np}{\sqrt{npq}} = \frac{X - \mu}{\sigma}, \text{ (say)}$$

where $\mu = np$ and $\sigma^2 = npq$, is given by

$$M_Z(t) = e^{-\mu t/\sigma} M_X(t/\sigma)$$

$$= e^{-npt/\sqrt{npq}} \cdot (q + p\, e^{t/\sqrt{npq}})^n \qquad \text{[From (**)]}$$

$$= \left[ e^{-pt/\sqrt{npq}} (q + p\, e^{t/\sqrt{npq}}) \right]^n$$

$$= \left[ qe^{-pt/\sqrt{npq}} + pe^{qt/\sqrt{npq}} \right]^n$$

$$= \left[ q \left\{ 1 - \frac{pt}{\sqrt{npq}} + \frac{p^2 t^2}{2npq} + 0' (n^{-3/2}) \right\} + p \left\{ 1 + \frac{qt}{\sqrt{npq}} + \frac{q^2 t^2}{2npq} + 0'' (n^{-3/2}) \right\} \right]^n$$

where $0' (n^{-3/2})$ and $0'' (n^{-3/2})$ involve terms containing $n^{3/2}$ and higher powers of $n$ in the denominator.

$$\therefore \; M_Z(t) = \left[ (q + p) + \frac{t^2 pq}{2npq} (p + q) + 0 (n^{-3/2}) \right]^n$$

$$= \left[ 1 + \frac{t^2}{2n} + 0 (n^{-3/2}) \right]^n$$

where $0 (n^{-3/2})$ involves terms with $n^{3/2}$ and higher powers of $n$ in the denominator.

$$\therefore \; \log M_Z(t) = n \log \left[ 1 + \frac{t^2}{2n} + 0 (n^{-3/2}) \right]$$

$$= n \left[ \left\{ \frac{t^2}{2n} + 0 (n^{-3/2}) \right\} - \frac{1}{2} \left\{ \frac{t^2}{2n} + 0 (n^{-3/2}) \right\}^2 + \ldots \right]$$

$$= \frac{t^2}{2} + 0''' (n^{-1/2})$$

where $0''' (n^{-1/2})$ involve terms with $n^{1/2}$ and higher powers of $n$ in the denominator. Proceeding to the limit as $n \to \infty$, we get

$$\lim_{n \to \infty} \log M_Z(t) = \frac{t^2}{2}$$

$$\Rightarrow \qquad \lim_{n \to \infty} M_Z(t) = \exp (t^2/2) \qquad \ldots(**)$$

**Interpretation.** (**) is the m.g.f. of standard normal variate [c.f. **Remark** to § 8·2·5]. Hence by uniqueness theorem of moment **generating functions,**

standard binomial variate tends to standard normal variate as $n \to \infty$. In other words, binomial distribution tends to normal distribution as $n \to \infty$.

**Example 7·20.** *A drunk performs a random walk over positions 0, ± 1, ± 2, ..., as follows. He starts at 0. He takes successive one unit steps, going to the right with probability p and to the left with probability (1 − p). His steps are independent. Let X denote his position after n steps. Find the distribution of (X + n)/2 and find E (X).* **(I.I.T. B.Tech., Dec. 1991)**

**Solution.** With the ith step of the drunk, let us associate a variable $X_i$ defined as follows :

$X_i = 1$, if he takes the step to the right

$= -1$ if he takes the step to the left

Then $X = X_1 + X_2 + \ldots + X_n$, gives the position of the drunkard after $n$ steps.

Define $\quad Y_i = (X_i + 1)/2$

Then $\quad Y_i = (1 + 1)/2 = 1$, with probability $p$

$= (-1 + 1)/2 = 0$, with probability $1 - p = q$, (say).

Since the $n$ steps of drunkard are independent, $Y_i$'s, $(i = 1, 2, \ldots n)$ are i.i.d. Bernoulli variates with parameter $p$.

Hence $\sum_{i=1}^{n} Y_i \sim B(n, p)$

$$\Rightarrow \sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \left(\frac{X_i + 1}{2}\right) = \frac{1}{2}\left[\sum_{i=1}^{n} X_i + n\right] = \frac{X + n}{2} \sim B(n, p)$$

where $X = \sum_{i=1}^{n} X_i$, is the position of the drunkard after $n$ steps.

Since $(X + n)/2 \sim B(n, p)$, we have

$$E\left[\frac{X + n}{2}\right] = np \Rightarrow \frac{1}{2} E(X + n) = np$$

$$\Rightarrow E(X) + n = 2np \Rightarrow E(X) = n(2p - 1)$$

**Example 7·21.** *Suppose that the r.v. X is uniformly distributed on (0,1)*

i.e., $f_X(x) = 1 ; 0 \le x \le 1.$ ...(*)

*Assume that the conditional distributional $Y | X = x$ has a binomial distribution with parameters n and p = x, i.e.,*

$$P(Y = y | X = x) = \binom{n}{y} x^y (1 - x)^{n-y} ; y = 0, 1, 2, \ldots, n \qquad (**)$$

*Find (a) E (Y)*

*(b) Find the distribution of Y.* **(Punjab P.C.S., 1990)**

**Solution.** (a) We are given that the conditional distribution of

$Y | X = x \sim B(n, x)$ ...(i)

$\therefore E(Y | X = x) = nx$ ...(ii)

We have :

$$E(Y) = E[E(Y|X)] = E[nX] = nE(X) \qquad \text{[On using (ii)]}$$

$$\text{Now } E(X) = \int_0^1 x f(x)\,dx = \int_0^1 x\,dx = \frac{1}{2}.$$

$$\therefore\ E(Y) = n \times (\tfrac{1}{2}) = \frac{1}{2} n$$

(b) We have : $f_{X,Y}(x, y) = f_X(x) \cdot f_{Y|X}(y \mid x)$

Since $X$ has (continuous) uniform distribution on (0,1) marginal distribution of $Y$ is given by.

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,dx = \int_0^1 f_{Y|X}(y \mid x) \cdot f_X(x)\,dx$$

$$= \int_0^1 {}^nC_y \cdot x^y (1 - x)^{n-y} \cdot 1 \cdot dx \qquad \text{[using (*) and (**)]}$$

$$= {}^nC_y \int_0^1 x^y (1 - x)^{n-y}\,dx$$

$$= {}^nC_y \cdot \beta(y + 1, n - y + 1) = \frac{n!}{y!\,(n - y)!} \cdot \frac{\Gamma(y + 1)\,\Gamma(n - y + 1)}{\Gamma(n + 2)}$$

$$= \frac{n!}{y!\,(n - y)!} \times \frac{y!\,(n - y)!}{(n + 1)!}$$

$$= \frac{1}{n + 1} \quad ; \quad y = 0, 1, 2, \ldots, n$$

Since $Y$ takes the values 0, 1, 2, ..., $n$ each with equal probability $1/(n + 1)$, $Y$ has discrete uniform distribution.

**Remark** We could find $E(Y)$ on using the distribution of $Y$ in (*b*).

$$E(Y) = \sum_{y=0}^{n} y\,p(y) = \frac{1}{n + 1} \sum_{y=0}^{n} y$$

$$= \frac{1}{n + 1}[0 + 1 + 2 + \ldots + n] = \frac{n}{2},$$

as in Part (a).

**Example 7·22.** *If $K(t)$ is the cumulative function about the origin of the Binomial Distribution of size n, show that*

$$\frac{d}{dt} K(t) = n \left\{1 + e^{-(z+t)}\right\}^{-1}, \text{ where } z = \log_e (p/q)$$

(b) *By expanding the R.H.S. in powers of t by Taylor's Theorem, show that*

$$\kappa_r = n \frac{d^{r-1} p}{dz^{r-1}}, \text{ where } \kappa_r \text{ is the } r\text{th cumulant.}$$

(c) *Hence or otherwise obtain the recurrence relation*

$$\kappa_{r+1} = pq \cdot \frac{d\kappa_r}{dp}, \; r > 1$$

[Baroda Univ. B.Sc. 1993; Delhi Univ. B.Sc. (Stat. Hons.) 1992]

(d) *Prove that* $\kappa_{r+1} = \dfrac{d\kappa_r}{dz}$, *where* $z = \log_e (p/q)$

**Solution.** For binomial distribution with parameters $n$ and $p$, we have

$$K(t) = \log M(t) = n \log (q + pe^t)$$

(a) $$\frac{d}{dt} K(t) = \frac{npe^t}{q + pe^t} = n\left(1 + \frac{q}{p} e^{-t}\right)^{-1}$$

if $z = \log_e (p/q) \Rightarrow (p/q) = e^z \Rightarrow (q/p) = e^{-z}$, then

$$\frac{d}{dt} K(t) = n[1 + e^{-(z+t)}]^{-1} \qquad \ldots(*)$$

(b) $$\kappa_r = \left[\frac{d^r}{dt^r} \kappa(t)\right]_{t=0} = \left[\frac{d^{r-1}}{dt^{r-1}} \cdot \frac{d}{dt} \kappa(t)\right]_{t=0}$$

$$= n\left[\frac{d^{r-1}}{dt^{r-1}} \left\{1 + e^{-(z+t)}\right\}^{-1}\right]_{t=0} = n\left[\frac{d^{r-1}}{dt^{r-1}}\left(\frac{e^{z+t}}{1 + e^{z+t}}\right)\right]_{t=0} \quad \ldots(**)$$

By summetry of the function $e^{z+t}/(1 + e^{z+t})$ in $t$ and $z$ we have

$$\frac{d}{dt}\left(\frac{e^{z+t}}{1 + e^{z+t}}\right) = \frac{d}{dz}\left(\frac{e^{z+t}}{1 + e^{z+t}}\right)$$

$$\Rightarrow \quad \frac{d^{r-1}}{dt^{r-1}}\left(\frac{e^{z+t}}{1 + e^{z+t}}\right) = \frac{d^{r-1}}{dz^{r-1}}\left(\frac{e^{z+t}}{1 + e^{z+t}}\right)$$

Substituting in (**), we get

$$\kappa_r = n\left[\frac{d^{r-1}}{dz^{r-1}}\left(\frac{e^{z+t}}{1 + e^{z+t}}\right)\right]_{t=0} = n \frac{d^{r-1}}{dz^{r-1}}\left(\frac{e^z}{1 + e^z}\right)$$

$$= n \frac{d^{r-1}}{dz^{r-1}} (1 + e^{-z})^{-1} = n \frac{d^{r-1}}{dz^{r-1}}\left(1 + \frac{q}{p}\right)^{-1}$$

$$= n \frac{d^{r-1} p}{dz^{r-1}} \qquad \ldots(***)$$

(c) $$\frac{d\kappa_r}{dp} = n \frac{d}{dp}\left(\frac{d^{r-1} p}{dz^{r-1}}\right) = n \frac{d}{dz}\left(\frac{d^{r-1} p}{dz^{r-1}}\right)\frac{dz}{dp}$$

$$= n \frac{d^r p}{dz^r} \cdot \frac{1}{pq} \qquad [\because z = \log_e (p/q)]$$

$$= \frac{1}{pq} \cdot \kappa_{r+1} \qquad \text{[From (***)]}$$

(d) $\dfrac{d\kappa_r}{dz} = \dfrac{d\kappa_r}{dp}\cdot\dfrac{dp}{dz} = \dfrac{d\kappa_r}{dp}\Big/\dfrac{dz}{dp} = \dfrac{d\kappa_r}{dp}\Big/\dfrac{1}{pq} = pq\cdot\dfrac{d\kappa_r}{dp}$

$\therefore \qquad \dfrac{d\kappa_r}{dz} = \kappa_{r+1}$ [c.f. part (c)]

**Example 7·23.** *If* $b(r\,;n,p) = \binom{n}{r} p^r q^{n-r}$ *is the binomial probability in the usual notation and if*

$$B(k\,;n,p) = P(X \le k) = \sum_{r=0}^{k} b(r\,;n,p),$$

*then prove that*

$$B(k\,;n,p) = (n-k)\binom{n}{k}\int_0^q t^{n-k-1}(1-t)^k\,dt;\ q = 1-p$$

**Solution.** $B(k\,;n,p) = \sum\limits_{r=0}^{k} b(r\,;n,p) = \sum\limits_{r=0}^{k}\binom{n}{r}p^r q^{n-r}$

Differentiating w.r. to $q$ and noting that $q = 1-p \Rightarrow \dfrac{dq}{dp} = -1$, we get:

$$\frac{d}{dq}\cdot B(k\,;n,p) = \sum_{r=0}^{k}\left[\binom{n}{r}\left\{r p^{r-1}(-1)\cdot q^{n-r} + p^r\cdot(n-r)q^{n-r-1}\right\}\right]$$

$$= \sum_{r=0}^{k}\left[\frac{n!\,(-r)}{r!\,(n-r)!}p^{r-1}q^{n-r} + \frac{n!\,(n-r)}{r!\,(n-r)!}p^r q^{n-r-1}\right]$$

$$= \sum_{r=0}^{k}\left[-\frac{n\,(n-1)!}{(r-1)!\,(n-r)!}p^{r-1}q^{n-r} + \frac{n\,(n-1)!}{r!\,(n-r-1)!}p^r q^{n-r-1}\right]$$

$$= \sum_{r=0}^{k}\left[n\cdot\binom{n-1}{r}p^r q^{n-r-1} - n\binom{n-1}{r-1}p^{r-1}q^{n-r}\right]$$

$$= \sum_{r=0}^{k}\left[n\left\{t_r - t_{r-1}\right\}\right] \qquad \ldots(**)$$

$$\text{where } t_r = \binom{n-1}{r}p^r q^{n-r-1} \qquad \ldots(***)$$

$$= n\left[\left(t_0 - t_{-1}\right) + \left(t_1 - t_0\right) + \left(t_2 - t_1\right) + \ldots + \left(t_k - t_{k-1}\right)\right]$$

$$= n\,t_k \qquad [\because t_{-1} = 0, \text{ From } (***)]$$

$$\therefore\ \frac{d}{dq}\cdot B(k,n,p) = n\binom{n-1}{k}p^k\cdot q^{n-k-1}, \quad p = 1-q$$

On integration, we get

$$B(k; n, p) = n \cdot \binom{n-1}{k} \int_0^q (1-u)^k \cdot u^{n-k-1}\, du.$$

But $n \cdot \binom{n-1}{k} = \dfrac{n \cdot (n-1)!}{k!\,(n-1-k)!} = \dfrac{n!\,(n-k)}{k!\,(n-k)!} = (n-k)\binom{n}{k}$

$$\therefore \quad B(k; n, p) = (n-k)\binom{n}{k} \int_0^q (1-u)^k \cdot u^{n-k-1}\, du$$

as desired.

**Remarks. 1.** We further get :

$$\beta(k+1, n-k) = \frac{\Gamma(k+1)\,\Gamma(n-k)}{\Gamma(n+1)} = \frac{k!\,(n-k-1)!}{n!}$$

$$\Rightarrow \frac{1}{\beta(k+1, n-k)} = \frac{n!}{k!\,(n-k-1)!} = (n-k)\binom{n}{k}$$

Hence the result may be written as :

$$B(k; n, p) = P(X \le k) = \frac{1}{\beta(k+1, n-k)} \int_0^q (1-u)^k u^{n-k-1}\, du$$

This result is of great practical utility. It enables us to represent the cumulative Binomial Probabilities (which are generally quite tedious and time consuming to compute) in terms of Incomplete Beta Functions which are tabulated in Karl Pearson's Tables of the Incomplete Beta Functions.

**2** Let us now work out the probability :

$$P(X \ge k) = \sum_{r=k}^{n} \binom{n}{r} p^r q^{n-r}$$

Differentiating w.r. to $p$, and proceeding similarly, we shall get :

$$\frac{d}{dp} P(X \ge k) = -n \sum_{r=k}^{n} \left(T_r - T_{r-1}\right) \qquad \text{(Try it)}$$

where $\qquad T_r = \binom{n-1}{r} p^r q^{n-r-1}, \qquad (T_n = 0)$

$$\therefore \frac{d}{dp} P(X \ge k) = n\,T_{k-1} = n\binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \qquad (\because q = 1-p)$$

On integration, we shall get :

$$P(X \ge k) = n\binom{n-1}{k-1} \int_0^p u^{k-1} (1-u)^{n-k}\, du$$

$$P(X \ge k) = \frac{1}{\beta(k, n-k+1)} \int_0^p u^{k-1} (1-u)^{n-k}\, du$$

This is quite an important result and should be committed to memory. We shall use it in *'Order Statistics'*.

This result can be stated as follows :

If $X \sim B(n, p)$ and $Y$ has Beta distribution with parameters $k$ and $n - k + 1$ (c.f. Chapter 8), then

$$P(Y \le p) = P(X \ge k) = 1 - P(X \le k - 1)$$

$$\Rightarrow \qquad F_Y(p) = 1 - F_X(k - 1)$$

## EXERCISE 7 (a)

1. (*a*) Describe the probability model from which the binomial distribution can be generated. Hence find the first four central moments.

(*b*) If $p$ is the probability of 'success' at a single trial, obtain the probability of $r$ 'successes' out of $n$ independent trails. Determine the mode of the resulting distribution.

2. (*a*) Define the binomial distribution with parameters $p$ and $n$, and give a situation in real life where the distribution is likely to be realized. Obtain the moment generating function of the binomial distribution and hence or otherwise obtain the mean, variance, skewness and kurtosis of the distribution.

(*b*) Obtain the Moment Generating Function of the Binomial Distribution. Derive from it the result that the sum of two binomial variates is a binomial variate if the variates are independent and have the same probability of success.

3. The mean and variance of a binomial variate $X$ with parameters $n$ and $p$ are 16 and 8. Find

(*i*) $P(X = 0)$, (*ii*) $P(X = 1)$, (*iii*) $P(X \ge 2)$.

4. For a Binomial distribution the mean is 6 and the standard deviation is $\sqrt{2}$. Write out all the terms of the distribution.

**Ans.** $n = 9,\ p = 2/3,\ q = 1/3;\ p(r) = (1/3)^9 \cdot \binom{9}{r} 2^r;\ r = 0, 1, 2, \ldots, 9$

5. (*a*) A perfect cube is thrown a large number of times in sets of 8. The occurrence of a 2 or 4 is called a success. In what proportion of the sets would you expect 3 successes.

**Ans** 27·31%

(*b*) In eight throws of a die, 5 or 6 is considered a success. Find the mean number of successes and the standard deviation. (Ans. 2·66, 1·33)

(*c*) A man tosses a fair coin 10 times. Find the probability that he will have

(*i*) heads on the first five tosses and tails on the next five tosses

(*ii*) heads on tosses 1, 3, 5, 7, 9 and tails on tosses 2, 4, 6, 8, 10.

(*iii*) 5 heads and 5 tails

(*iv*) at least 5 heads

(*v*) not more than 5 heads. **[Madras Univ. B.Sc. (Main Stat) Nov. 1991]**

**Ans.** (*i*) $(1/2)^{10}$, (*ii*) $(1/2)^{10}$, (*iii*) ${}^{10}C_5 (1/2)^{10}$

(*iv*) $\sum_{x=5}^{10} {}^{10}C_x \left(\frac{1}{2}\right)^{10}$ (*v*) $\sum_{x=0}^{5} {}^{10}C_x \left(\frac{1}{2}\right)^{10}$

6. (*a*) In 256 sets of twelve tosses of a fair coin, in how many cases may one expect eight heads and four tails ?

**(Ans. 31)** **(Delhi Univ. B.Sc. Oct. 1992)**

(*b*) In 100 sets of ten tosses of an unbaised coin, in how many cases should we expect

(*i*) Seven heads and three tails, (*ii*) at least seven heads ?

Ans. (*i*) 12, (*ii*) 17

7. (*a*) During war 1 ship out of 9 was sunk on an average in making a certain voyage. What was the probability that exactly 3 out of a convoy of 6 ships would arrive safely ? (Madras Univ. B.Sc., 1992)

Ans. $^6C_3\ (8/9)^3\ (1/9)^3$

(*b*) In the long run 3 vessels out of every 100 are sunk. If 10 vessels are out, what is the probability that

(*i*) exactly 6 will arrive safely, and

(*ii*) at least 6 will arrive safely ?

**Hint.** The probability '$p$' that a vessel will arrive safely is

$$P = 97/100 = 0\cdot 97 \text{ and } q = 0\cdot 03$$

The probability that out of 10 vessels, $x$ vessels will arrive safely is

$$p(x) = {}^{10}C_x p^x q^{10-x} = {}^{10}C_x (0\cdot 97)^x (0\cdot 03)^{10-x}$$

(*i*) Required probability = $p(6) = {}^{10}C_6 (0\cdot 97)^6 (0\cdot 03)^x$.

(*ii*) Required probability= $P\ (X \geq 6)$

8. (*a*) A student takes a true-false examination consisting of 10 questions. He is completely unprepared so he plans to guess each answer. The guesses are to be made at random. For example, he may toss a fair coin and use the outcome to determine his guess.

(*i*) Compute the probability that he guesses correctly at least five times.

(*ii*) Compute the probability that he guesses correctly at least 9 times.

(*iii*) What is the smallest $n$ that the probability of guessing at least $n$ correct answers is less than 1/2. (Dibrugarh Univ. M.A., 1993)

Ans. (*i*) 319/512 ; (*ii*) 11/1024 ; (*iii*) 6.

(*b*) A multiple choice test consists of 8 questions and 3 answers to each question, of which only one is correct. If a student answers each question by rolling a balanced die and checking the first answer if he gets 1 or 2, the second answer if he gets 3 or 4, and the third answer if he gets 5 or 6, find the probability of getting:

(*i*) exactly 3 correct answers,

(*ii*) no correct answer,

(*iii*) at least 6 correct answers. [Gauhati Univ. M.A. (Econ.), 1993]

9. (*a*) The incidence of occupational disease in an industry is such that the workers have a 20% chance of suffering from it. What is the probability that out of six workers chosen at random, four or more will suffer from the disease.

Ans. 52/3125

(*b*). (*a*) In a binomial distribution consisting of 5 independent trails, probabilities of 1 and 2 successes are 0·4096 and 0·2048 respectively. Find the parameter $p$ of the distribution. (Ans. 0·2)

10. (*a*) With the usual notations, find $p$ for a binomial random variable $X$ if $n = 6$ and if $9P(X = 4) = P(X = 2)$. (Ans. 0·25)

**(Mysore Univ. B.Sc. April 1992)**

(*b*) $X$ is a random variable following binomial distribution with mean 2·4 and variance 1·44. Find $P(X \geq 5)$, $P(1 < X \leq 4)$.

11. (*a*) In a certain town 20% of the population is literate, and assume that 200 investigators take a sample of ten individuals each to see whether they are literate. How many investigators would you expect to report that three people or less are literates in the sample? **(Shivaji Univ. B.Sc., Oct. 1992)**

(*b*) A lot contains 1 per cent of defective items. What should be the number ($n$) of items in a random sample so that the probability of finding at least one defective in it, is at least 0·95? (Ans. 68)

12. (*a*) If on the average rain falls on ten days in every thirty days, find the probability

(*i*) that rain falls on at least three days of a given week,

(*ii*) that first three days of a given week will be dry and the remaining wet.

Ans. (*i*) $\sum_{x=3}^{7} {}^7C_x (1/3)^x (2/3)^{7-x}$, (*ii*) $(2/3)^3 \cdot (1/3)^4$.

(*b*) Suppose that weather records show that on the average 5 out of 31 days in October are rainy days. Assuming a binomial distribution with each day of October as an independent trial, find the probability that the next October will have at most three rainy days.

Ans. 0·2403

13. The probability of a man hitting a target is 1/4. (*i*) If he fires 7 times, what is the probability $p$ of his hitting the target at least twice? (*ii*) How many times must he fire so that the probability of his hitting the target at least once is greater than 2/3? [ Ans. (*i*) 4547/8192, (*ii*) 4 ]

**Hint.** (*ii*) $p = \frac{1}{4}$, $q = \frac{3}{4}$ · We want $n$ such that

$$1 - q^n > \frac{2}{3} \Rightarrow q^n < \frac{1}{3} \Rightarrow \left(\frac{3}{4}\right)^n < \frac{1}{3} \Rightarrow n = 4$$

14. (*a*) The probability of a man hitting a target is 1/3. How many times must be fire so that the probability of hitting the target at least once is more than 90%. Ans. 6. **(Shivaji Univ. B.Sc., 1991)**

(*b*) Eight mice are selected at random and they are divided into two groups of 4 each. Each mouse in group $A$ is given a dose of certain poison '*a*' which is expected to kill one in four; each mouse in group $B$ is given a dose of certain poison '*b*' which is expected to kill one or two. Show that nevertheless, there may be fewer deaths in group $A$ and find the probability of this happening.

Ans. 525/4096

15 (*a*) A card is drawn and replaced in an ordinary deck of 52 cards. How many times must a card be drawn so that (*i*) there is at least an even chance of drawing a heart, (*ii*) the probability of drawing a heart is greater than 3/4?

Ans. (*i*) 3, (*ii*) 5

(*b*) Five coins are tossed. What is the variance of the number of heads per toss of the five coins:

(*i*) if each coins is unbiased,

(*ii*) if the probability of a head appearing is 0· 75 for each coin, and

(*iii*) if four coins are unbaised and for the fifth the probability of a head appearing is 0· 75 ?

**Hint** (*iii*) Use generating function. [See Ex. 7· 17 (*iii*)]

**16.** An owner of a small hotel with five rooms is considering buying television sets to rent to room occupants. He expects that about half of his customers would be willing to rent sets, and finally he buys three sets. Assuming 100% occupancy at all times :

(*i*) . What fraction of the evenings will there be more request than T.V. sets?

(*ii*) What is the probability that a customer who requests a television set will receive one?

(*iii*) If the owner's cost per set per day is $C$, what rent $R$ must he charge in order to break even (neither gain nor lose) in the long run ?

**Hint.** (*i*) Let the random variable $X$ denote the daily number of requests. Then required probability is

$$P(X \geq 4) = P(X = 4) + P(X = 5) = \binom{5}{4}\left(\frac{1}{2}\right)^5 + \binom{5}{5}\left(\frac{1}{2}\right)^5$$

(*ii*) The customer can get a T.V. in the following mutually exclusive ways,

(*a*) There are no other requests that night.

(*b*) There is one other request.

(*c*) There are two other requests.

(*d*) There are three other requests and his request precedes at least one of them.

(*e*) There are four other requests, and his request preceedes at least two of them.

The probability of the desired event

$$= (0\cdot 5)^4 \left\{1 + {}^4C_1 + {}^4C_2 + \frac{3}{4}\cdot {}^4C_3 + \frac{3}{5}\,{}^4C_4\right\}$$

(*iii*) Mean revenue

$$= (0\cdot5)^5 \cdot 0 + {}^5C_1\,(0\cdot5)^5 R + {}^5C_2\,(0\cdot5)^5\, 2R + \left\{{}^5C_3\,(0\cdot5)^5 + {}^5C_4\,(0\cdot5)^5 + {}^5C_5\,(0\cdot5)^5\right\} 3R$$

$$= \frac{73}{32} R$$

The break-even rental is the value of $R$ for which

$$\frac{73}{32} R = 3C \quad \Rightarrow \quad R = 1\cdot 315\, C$$

**17.** A manufacturer claims that at most 10 per cent of his product is defective. To test this claim, 18 units are inspected and his claim is accepted if among these 18 units, at most 2 are defective. Find the probability that the manufacturer's claim will be accepted if the actual probability that a unit is defective is

(*a*) 0· 05 (*b*) 0· 10 (*c*) 0· 15 and (*d*) 0· 20.

**Ans.** (*a*) 0·9410 (*b*) 0·9326 (*c*) 0·4445 (*d*) 0·2715

**18.** (*a*) A set of 8 symmetrical coins was tossed 256 times and the frequencies of throws observed were as follows :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Number of heads : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Frequency of throws: | 2 | 6 | 24 | 63 | 64 | 50 | 36 | 10 | 1 |

Fit a binomial distribution and find mean and standard deviation of fitted distribution.

(*b*) A set of 6 similar coins is tossed 640 times with the following results:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Number of heads : | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| Frequency : | 7 | 64 | 140 | 210 | 132 | 75 | 12 |

Calculate the binomial frequencies on the assumption that the coins are symmetrical.

**19.** (*a*) The following data due to Weldon shows the results of throwing 12 dice 4096 times, a throw of 4, 5 or 6 being called a success ($x$).

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Total |
| $f$ : | — | 7 | 60 | 198 | 430 | 731 | 948 | 847 | 536 | 257 | 71 | 11 | — | 4096 |

Fit the binomial distribution and calculate the expected frequencies. Compare the actual mean and S.D. with those of the expected ones for the distribution.

**Ans.** Expected freq. : 1, 12, 66, 220 495 792, 924, 792, 495, 220, 66, 12, 0; mean = 6, variance = 1·71.

(*b*) In 103 litters of 4 mice, the number of litters which contained 0, 1, 2, 3, 4 females are recorded below :

| Number of female mice | 0 | 1 | 2 | 3 | 4 | Total |
|---|---|---|---|---|---|---|
| Number of litters | 8 | 32 | 34 | 24 | 5 | 103 |

(*i*) If the chance of obtaining a female in a single trial is assumed constant, estimate the constant but unknown probability.

(*ii*) If the size of the litter 4 had not been given, how could it be estimated from the data ?

**20.** $X$ is random variable distributed according to the Binomial law :

$$b(x\,;n,p) = \binom{n}{x} p^x q^{n-x}\,;\; x = 0, 1, 2, ..., n$$

Obtain the recurrence formula :

$$b(x+1\,;n,p) = \frac{n-x}{x+1}\cdot\frac{p}{q}\cdot b(x\,;n,p)$$

Use this as a reduction formula and get the theoretical frequencies when an unbaised coin is tossed 8 times and the experiment is repeated 256 times.

**(Madras Univ. B. Sc. April 1992)**

**21.** (*a*) By differentiating the following identity with respect to $p$ and then multiplying by $p$,

$$\sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (q+p)^n, q = 1-p$$

prove that $\mu_1' = np$ and $\mu_2 = npq$.

**22.** (*a*) Let $X \sim b(x; n, p)$ and $r$ be a non-negative integer. If the $r$th moment about the origin is denoted by $\mu_r' = E(X^r)$, prove that

$$\mu_{r+1}' = np\,\mu_r' + p(1-p)\frac{d\mu_r'}{dp}$$

**[Delhi Univ. B.Sc. (Hons. Subs.), 1993, '88]**

(*b*) Show that for the binomial distribution $B(n, p)$,

$$\mu_{r+1} = pq\left(nr\,\mu_{r-1} + \frac{d}{dp}\mu_r\right), \quad p + q = 1,$$

where symbols have their usual meanings.

**[Delhi Univ. B.Sc. (Stat. Hons), 1989]**

(c) If $X \sim B(n, p)$, obtain the recurrence relation for its central moments and hence find values of $\beta_1$ and $\beta_2$.

**[Calcutta Univ. B.Sc. (Hons.), 1992]**

**23.** (*a*) The following results were obtained when 100 batches of seeds were allowed to germinate on damp filter paper in a laboratory :

$$\beta_1 = \frac{1}{15} \text{ and } \beta_2 = \frac{89}{30}$$

Determine the binomial distribution and calculate the frequency for $X = 8$, considering $p > q$.

**Hint.** We have $\beta_1 = \dfrac{(q-p)^2}{npq} = \dfrac{1}{15}$ ...(*i*)

and $$\beta_2 = 3 + \frac{1-6pq}{npq} = \frac{89}{30} \quad \text{...(ii)}$$

From (*i*) and (*ii*), we can find the value of $n, p$ and $q$

(*b*) Between a Binomial distribution with $n = 5$ and $p = \frac{1}{2}$ and a distribution with frequency function

$$f(x) = 6x(1-x),\ 0 \le x \le 1;$$

determine which is more skewed.

**24.** (*a*) $x = r$ is the unique mode of Binomial Distribution having mean $np$ and variance $np(1-p)$. Show that

$$(n+1)p - 1 < r < (n+1)p$$

Find the mode of the binomial distribution with $p = \frac{1}{2}$ and $n = 7$.

**[Delhi Univ. B.Sc. (Stat. Hons.) 1991, '84]**

**Ans.** 4, 3 (Bimodal).

(*b*) Show that if $np$ be a whole number, the mean of the binomial distribution coincides with the greatest term.

(c) Compute the mode of a binomial distribution $b\,(7, \frac{1}{2})$.

**[Delhi Univ. B.Sc. (Maths. Hons.), 1989]**

**Ans.** 1, 2 (Bimodal).

(d) Define Bernoulli trials and state the binomial law of probability. Find the bounds for the most probable number of successes in a sequence of $n$ Bernoulli trials.

One workers can manufacture 120 articles during a shift, another worker 140 articles, the probabilities of the articles being of a high quality are 0·94 and 0·80 respectively. Determine the most probable number of high quality articles manufactured by each worker. **[Calcutta Univ. B.Sc. (Maths. Hons.), 1988]**

**25.** Show that if two symmetrical binomial distributions ($p = q = \frac{1}{2}$) of degree $n$ (and of the same number of observations) are so superimposed that the $r$th term of one coincides with the $(r+1)$th term of the other, the distribution formed by adding superimposed terms is a symmetrical binomial of degree $(n+1)$. **[Bhagalpur Univ. B.Sc., 1993]**

**26.** (a) Let $X$ denote a binomially distributed random variable. Show that

$$E\left(\frac{X-np}{\sqrt{npq}}\right) = 0,\; E\left(\frac{X-np}{\sqrt{npq}}\right)^2 = 1, \text{ and}$$

$$E\left[\exp.\left\{t\left(\frac{X-np}{\sqrt{npq}}\right)\right\}\right] = \left[(1-p)\exp\left\{-t\sqrt{\left(\frac{p}{nq}\right)}\right\} + p\exp\left\{t\sqrt{\left(\frac{q}{np}\right)}\right\}\right]$$

(b) Obtain the characteristic function of the standard binomial variate $(X-np)/\sqrt{npq}$, where $X$ is the number of successes obtained in $n$ independent trials, each with constant probability $p$ of success, $q = 1-p$. Obtain the limit of this function as $n \to \infty$. **[Delhi Univ. B.Sc.(Maths. Hons.), 1991]**

(c) If $X \sim B\,(n, p)$, prove that

$$\kappa_{r+1} = pq \cdot \frac{d}{dp}(\kappa_r),$$

where $\kappa_r$ is the $r$th cumulant.

Hence deduce the values of $\kappa_2$ and $\kappa_3$.

**[Delhi Univ. B.Sc. (Stat. Hons.), 1991, '87]**

**27.** (a) If $X$ and $Y$ are two independent identically distributed binomial variates, obtain the probability that the absolute difference $|X - Y|$ equals a given value, say $r$.

(b) (i) If $X$ and $Y$ are independent binomial variates, with parameters $p_1$ and $p_2$ and indices $n_1$ and $n_2$ respectively, obtain the probability that $X + Y$ equals '$r$'.

(ii) In the above if $p_1 = p_2$, what is the distribution of $X + Y$ ?

**[Poona Univ. B.Sc., 1988]**

(*c*) If $X$ and $Y$ are two independent binomial variates with parameters $n_1 = 6$, $p = 1/2$ and $n_2 = 4$, $p = 1/2$ respectively, evaluate,

(*i*) $P(X + Y = r)$, (*ii*) $P(X + Y \geq 3)$

(Gujarat Univ. B. Sc. Oct. 1992)

**Hint** $X + Y \sim B(6 + 4, 1/2) = B(10, 1/2)$

**Ans.** (*i*) $P(X + Y = r) = p(r) = {}^{10}C_r\,(1/2)^r;\ r = 0, 1, \ldots, 10$

(*ii*) $P(X + Y \geq 3) = 1 - [p(0) + p(1) + p(2)] = 0\cdot945$

(*d*) If $X$ and $Y$ are two independent binomial variates with parameters $(n_1 = 3, p = 0\cdot4)$ and $(n_2 = 4, p = 0\cdot4)$ respectively, find:

(*i*) $P(X = Y)$, (*ii*) $P(X + Y \leq 2)$, (*iii*) $P(X = 3 \mid X + Y = 4)$

**Hint.** $X + Y \sim B(3 + 4, 0\cdot4) = B(7, 0\cdot4)$

(*i*) $P(X = Y) = \sum_{r=0}^{3} P(X = r \cap Y = r) = \sum_{r=0}^{3} P(X = r)\,P(Y = r) = 0\cdot2871$

(*ii*) $P(X + Y \leq 2) = \sum_{r=0}^{2} \binom{7}{r} (0\cdot4)^r (0\cdot6)^{7-r} = 0\cdot420$

(*iii*) $P(X = 3 \mid X + Y = 4) = \dfrac{P(X = 3 \cap X + Y = 4)}{P(X + Y = 4)} = \dfrac{P(X = 3 \cap Y = 1)}{P(X + Y = 4)} = 0\cdot1141$

**28.** (*a*) Obtain the moment generating function of Binomial distribution with $n = 7$ and $p = 0\cdot6$. Find the first three moments of the distribution.

[Poona Univ. B. Sc. 1992)

Ans. $(0\cdot4 + 0\cdot6\,e^t)^7$ : mean $= 4\cdot2$, $\mu_2 = 1\cdot68$, $\mu_3 = -0\cdot336$.

(*b*) Suppose that the m.g.f. of a random variable $X$ is of the form

$$M_X(t) = (0\cdot4\,e^t + 0\cdot6)^8$$

What is the m.g.f. of the random variable $Y = 3X + 2$? Evaluate $E(X)$.

Ans. $E(X) = 3\cdot2$, $M_Y(t) = e^{2t}\,(0\cdot6 + 0\cdot4\,e^{3t})^8$

(*c*) Obtain the moment generating function of the binomial distribution. Hence or otherwise obtain the mean, variance and skewness of the distribution.

**29.** Show that the factorial moment generating function $w(t)$ of the binomial distribution $b(x; n, p)$ is $(1 + pt)^n$. Hence or otherwise show that

$$\mu_{(r)}' = n^{(r)}\,p^r$$

**Hint.** Factorial moment generating function $\omega(t)$ is defined as

$$\omega(t) = E(1 + t)^X = \sum_x (1 + t)^x p(x) = \sum_x {}^nC_x \left\{p(1 + t)\right\}^x q^{n-x}$$

$$\mu_{(r)}' = \text{co efficient of } \frac{t^r}{r!} \text{ in } \omega(t) = {}^nC_r\, r!\, p^r = n^{(r)} p^r$$

**30.** Show that

(*i*) $$b(n, p; k) = b(n, 1 - p; n - k)$$

(*ii*) $$\sum_{k=r}^{n} b(n, p; k) = 1 - \sum_{k=n-r+1}^{n} b(n, 1 - p; k)$$

(*iii*) $$b(n + 1, p; k) = p\,.\,b(n, p; k - 1) + q\,.\,b(n, p; k)$$

**Hint.** (*i*) $b(n, 1-p; n-k) = \binom{n}{n-k} (1-p)^{n-k} p^{n-(n-k)}$

(*ii*) $\sum_{k=r}^{n} b(n, p; k) = \sum_{k=r}^{n} b(n, 1-p; n-k) = \sum_{k=0}^{n-r} b(n, 1-p; k)$

**31.** For a binomial distribution, let

$$F_n(y) = \sum_{x=0}^{y} \binom{n}{x} p^x q^{n-x},$$

where $q = 1-p$,

prove that

(*i*) $F_{n+1}(y) = p\,F_n(y-1) + q\,F_n(y)$

(*ii*) $\text{Cov}(X, n-X) = -npq$ **(Bombay Univ. B.Sc., April 1990)**

**32.** (*a*) Random variable $X$ follows binomial distribution with parameters $n = 40$ and $p = \frac{1}{4}$. Use Chebychev's inequality to find bounds for

(*i*) $P[\,|X-10| < 8]$ ; (*ii*) $P[\,|X-10| > 10]$

Compare these values with the actual values (**Hint** : Use Normal approximation for the Binomial). **(Madras Univ. B.Sc. (Main Stat.), 1988)**

**Ans.** (*i*) 113/128 (lower bound), (*ii*) 0.075 (upper bound).

(*b*) $X$ follows binomial distribution with $n = 40$, $p = \frac{1}{2}$. Use Chebychev's lemma to

(*i*) find $k$ such that

$P\{\,|X-20| > 10k\} \leq 0{\cdot}25$, and

(*ii*) obtain a lower limit for $P\{|X-20| \leq 5\}$.

**[Delhi Univ. B.Sc. (Maths. Hons.), 1984]**

**Ans.** (*i*) $2\sqrt{10}$, (*ii*) 3/5

(*c*) How many trials must be made of an event with binomial probability of success $\frac{1}{2}$ in each trial, in order to be assured with probability of at least 0·9 that the relative frequency of success will be between 0·48 and 0·52 ? (**Ans.** 6250)

**Hint.** Use Chebychev's Inequality.

**33.** (*a*) Show that if a coin is tossed $n$ times, the probability of not more than $k$ heads is :

$$\left[\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{k}\right] \left(\frac{1}{2}\right)^n$$

**[South Gujarat Univ. B.Sc., 1988]**

(*b*) If $X$ has binomial distribution with paramctes $n$ and $p$, then prove that $P[X \text{ is even}] = \frac{1}{2}[1 + (q-p)^n]$. **[Delhi Univ. B.Sc. (Stat. Hons.), 1988]**

**34.** If the probability of hitting a target is 1/5 and if 10 shots are fired, what is the conditional probability of the target being hit at least twice assuming that at least one hit is already scored ?

**[Nagpur Univ. B.Sc., 1988, '93]**

**Hint.** Let $X$ donote the number of times a target is hit when 10 shots are fired. Then $X \sim B(10, 0\cdot 2)$. The required probability is :

$$P(X \geq 2 \mid X \geq 1) = \frac{P[(X \geq 2) \cap (X \geq 1)]}{P(X \geq 1)} = \frac{P(X \geq 2)}{P(X \geq 1)}$$

$$= \frac{1 - [P(X = 0) + P(X = 1)]}{1 - [P(X = 0)]} = \frac{0\cdot 625}{0\cdot 893} = 0\cdot 6999$$

35. (*a*) Let $X$ be a $B(2, p)$ and $Y$ be a $B(4, p)$. If $P(X \geq 1) = 5/9$, find $P(Y \geq 1)$ **[Kerala Univ. B. Sc., 1989]**

**Hint.** $P(X \geq 1) = 1 - P(X = 0) = 1 - q^2 = 5/9 \Rightarrow q = 2/3, p = 1/3.$

$P(Y \geq 1) = 1 - P(Y = 0) = 1 - q^4 = 65/81.$

36. Let $B$ denote the number of boys in a family with five children. If $p$ denotes the probability that a boy is there in a family, find the least value of $p$ such that

$P(B = 0) > P(B = 1)$ **(Shivaji Univ. B. Sc., 1990)**

Ans. $q^5 > 5pq^4 \Rightarrow q > 5p \Rightarrow p < \frac{1}{6}.$

37. (*a*) Suppose $X \sim B(n, p)$. with $E(X) = 5$, $\text{Var}(X) = 4$. Find $n$ and $p$. (Ans. $n = 25$, $p = 1/5$)

(*b*) Let $X \sim B(n, p)$. For what $p$ is variance $(X)$ maximised if we assume $n$ is fixed.

Ans. $\text{Var}\, X = npq = n(p - p^2) = f(p)$, (say) ; $f'(p) = 0$, $f''(p) < 0$ ; $p = 1/2 = q$

38. (*a*) $X \sim B(n = 100, p = 0\cdot 1)$. Find $P(X \leq \mu_x - 3\sigma_x)$

Ans. $\mu = 10$, $\sigma = 3$, $P(X \leq \mu_x - 3\sigma_x) = P(X \leq 1) = 10\cdot 9 \times (0\cdot 9)^{99}$

(*b*) If $X \sim B(25, 0\cdot 2)$, find $P(X < \mu_x - 2\sigma_x)$

**[Delhi Univ. B.A. (Stat. Hons.) Spl. Course 1989]**

39. For one half of $n$ events, the chance of success is $p$, and the chance of failure is $q$, whilst for the other half the chance of success is $q$, and the chance of failure is $p$. Show that the S.D. of the number of successes is the same as if the chance of success were $p$ in all the cases i.e. $\sqrt{npq}$, but that the mean of the number of successes is $n/2$ and not $np$. **(Delhi Univ. B.A. 1992)**

**Hint.** $X \sim B(n/2, p)$ and $Y \sim B(n/2, q)$ are independent. Let $Z = X + Y$. Now prove that $\text{Var}(Z) = npq$ and $E(Z) = n/2$.

40. The discrete density of $X$ is given by $f_X(x) = x/3$, for $x = 1, 2$ and $f_{Y|X}(y \mid x)$ is binomial with parameters $x$ and $\frac{1}{2}$ i.e.,

$$F_{Y|X}(y \mid x) = P(Y = y \mid X = x) = \binom{x}{y} \cdot \left(\tfrac{1}{2}\right)^x ;$$

for $y = 0, 1, \ldots, x$ and $x = 1, 2$.

(*a*) Find $E(X)$ and $\text{Var}(X)$ ; (*b*) Find $E(Y)$

(c) Find the joint distribution of $X$ and $Y$.

**Hint.** **Proceed as in Example 7·21.**

Ans. (a) $E(X) = 5/3$, $\text{Var}(X) = 2/9$, (b) $E(Y) = 5/6$.

(c) $f(x,y) = \binom{x}{y} \cdot \left(\frac{x}{3}\right) \cdot \left(\frac{1}{2}\right)^x$ ; $n = 1, 2, ;$ $y = 0, 1, ..., x$.

41. Two dice are thrown $n$ times. Let $X$ denote the number of throws in which the number on the first dice exceeds the number on the second dice. What is the distribution of $X$?

Ans. $X \sim B(n, p = 15/36)$

**Hint.** $p$ is the probability that the number on the first dice exceeds the number on the second dice in a throw of two dice.

42. Let $X_1 \sim B(n, p_1)$ and $X_2 \sim B(n, p_2)$.

If $p_1 < p_2$, prove that:

$$P(X_1 \le k) \ge P(X_2 \le k) \text{ for } k = 0, 1, ..., n.$$

**Hint.** Use Example 7·23.

43. If $X \sim B(n, p)$, show that

$$P(X \le k) = \lambda \int_{p/q}^{\infty} \frac{y^k}{(1 + y)^{n+1}} \, dy$$

where $$\lambda^{-1} = \int_0^{\infty} \frac{y^k}{(1 + y)^{n+1}} \, dy = \beta(k + 1, n - k)$$

**Hint.** $\frac{d}{dq} P(X \le k) = n \binom{n-1}{k} \cdot p^k \cdot q^{n-k-1} = A_k$, (*say*)

[See Example 7·23]

Find $$\frac{d}{dq}(\text{RHS}) = \lambda \cdot \frac{d}{dq}\left(\int_{p/q}^{\infty} \frac{y^k}{(1 + y)^{n+1}} \, dy\right) = \lambda \frac{(p/q)^k}{[1 + (p/q)]^{n+1}} \left(\frac{1}{q^2}\right)$$

$$= \frac{1}{\beta(k + 1, n - k)} \cdot p^k \cdot q^{n-k-1} = A_k$$

(On simplification)

44. If $X$ $B(n, p)$ and $Y$ has beta distribution with parameters $k$ and $n - k + 1$, (See Chapter 8), then prove that

$P(Y \le p) = P(X \ge k)$ *i.e.*, $F_Y(p) = 1 - F_X(k - 1)$

45. If a fair coin is tossed an even number $2n$ times, show that the probability of obtaining more heads than tails is

$$\frac{1}{2}\left\{1 - {}^{2n}C_n \left(\frac{1}{2}\right)^{2n}\right\}$$

**Hint.** $X$ : No. of heads; $Y$ = No. of tails; No. of trials = $2n$

$P(X > Y) + P(X < Y) + P(X = Y) = 1$

$$\Rightarrow \quad 2P(X > Y) = 1 - P(X = Y)$$

[$\because$ By symmetry, $p = q = \frac{1}{2} \Rightarrow P(X > Y) = P(X < Y)$]

$$= 1 - {}^{2n}C_n p^n \cdot q^n = 1 - {}^{2n}C_n \left(\tfrac{1}{2}\right)^{2n}$$

$$\Rightarrow \qquad P(X > Y) = \frac{1}{2}\left[1 - {}^{2n}C_n \left(\tfrac{1}{2}\right)^{2n}\right]$$

**7·3·0. Poisson Distribution (as a limiting case of Binomial Distribution).** Poisson distribution was discovered by the French mathematician and physicist Simeon Denis Poisson (1781—1840) who published it in 1837. Poisson distribution is a limiting case of the binomial distribution under the following conditions:

(*i*) $n$, the number of trials is indefinitely large, *i.e.*, $n \to \infty$.

(*ii*) $p$, the constant probability of success for each trial is indefinitely small, *i.e.*, $p \to 0$.

(*iii*) $np = \lambda$, (say), is finite. Thus $p = \lambda/n$, $q = 1 - \lambda/n$, where $\lambda$ is a positive real number.

The probability of $x$ successes in a series of $n$ independent trials is

$$b(x\,;n,p) = \binom{n}{x} p^x q^{n-x}\,;\, x = 0, 1, 2, \ldots, n \qquad \ldots(*)$$

We want the limiting form of (*) under the above conditions. Hence

$$\lim_{n\to\infty} b(x\,;n,p) = \lim_{n\to\infty} \frac{n\,!}{x\,!\,(n-x)\,!}\left(\frac{\lambda}{n}\right)^x \cdot \left[1 - \frac{\lambda}{n}\right]^{n-x}$$

Using Stirling's approximation for $n\,!$ as $n \to \infty$ *viz.*,

$$\lim_{n\to\infty} n\,! \approx \sqrt{2\pi}\, e^{-n}\, n^{n+(1/2)}, \text{ we get}$$

$$\lim_{n\to\infty} b(x;n,p) = \lim_{n\to\infty}\left[\frac{\sqrt{2\pi}\,e^{-n}\cdot n^{n+(1/2)}}{x\,!\sqrt{2\pi}\,e^{-(n-x)}\cdot(n-x)^{n-x+(1/2)}}\right]\left(\frac{\lambda}{n}\right)^x\left[1-\frac{\lambda}{n}\right]^{n-x}$$

$$= \frac{\lambda^x}{e^x\cdot x\,!}\cdot\lim_{n\to\infty}\frac{n^{n-x+(1/2)}}{(n-x)^{n-x+(1/2)}}\cdot\left[1-\frac{\lambda}{n}\right]^{n-x}$$

$$= \frac{\lambda^x}{e^x\,x\,!}\lim_{n\to\infty}\frac{\left(1-\dfrac{\lambda}{n}\right)^{n-x}}{\left(1-\dfrac{x}{n}\right)^{n-x+(1/2)}}$$

$$= \frac{\lambda^x}{e^x\,x\,!}\cdot\frac{\lim\limits_{n\to\infty}\left[1-\dfrac{\lambda}{n}\right]^n\cdot\lim\limits_{n\to\infty}\left(1-\dfrac{\lambda}{n}\right)^{-x}}{\lim\limits_{n\to\infty}\left[1-\dfrac{x}{n}\right]^n\lim\limits_{n\to\infty}\left[1-\dfrac{x}{n}\right]^{-x+(1/2)}}$$

But we know that

and
$$\left.\begin{array}{l} \lim\limits_{n \to \infty} \left(1 - \dfrac{\lambda}{n}\right)^{n} = e^{-\lambda}, \\[2ex] \lim\limits_{n \to \infty} \left[1 - \dfrac{\lambda}{n}\right]^{\alpha} = 1,\ \alpha \text{ is not a function of } n \end{array}\right\} \quad ...(**)$$

Therefore
$$\lim_{n \to \infty} b(x ; n, p) = \frac{\lambda^x}{e^x \cdot x!} \cdot \frac{e^{-\lambda} \cdot 1}{e^{-x} \cdot 1} = \frac{e^{-\lambda} \cdot \lambda^x}{x!} ; x = 0, 1, 2, ..., \infty ;$$

[Using (**)]

which is the required probability function of the Poisson distribution. 'λ' is known as the parameter of Poisson distribution.

**Aliter.** Poisson distribution can also be derived without using Stirling's approximation as follows :

$$b(x ; n, p) = \binom{n}{x} p^x (1 - p)^{n-x} = \binom{n}{x} \left[\frac{p}{1-p}\right]^x (1 - p)^n$$

$$= \frac{n(n-1)(n-2)\dots(n-x+1)}{x!} \cdot \frac{\left(\dfrac{\lambda}{n}\right)^x}{\left[1 - \dfrac{\lambda}{n}\right]^x} \left[1 - \frac{\lambda}{n}\right]^n$$

$$= \frac{\left[1 - \dfrac{1}{n}\right]\left[1 - \dfrac{2}{n}\right]\dots\left[1 - \dfrac{x-1}{n}\right]}{x!\left[1 - \dfrac{\lambda}{n}\right]^x} \lambda^x \left[1 - \frac{\lambda}{n}\right]^n$$

$$\therefore \lim_{n \to \infty} b(x ; n, p) = \frac{e^{-\lambda}\lambda^x}{x!} ; x = 0, 1, 2, \dots$$ [From (**)]

**Definition.** *A random variable X is said to follow a Poisson distribution if it assumes only non-negative values and its probability mass function is given by*

$$p(x, \lambda) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} ; x = 0, 1, 2, \dots ; \lambda > 0$$
$$= 0, \textit{ otherwise} \qquad ...(7\cdot 14)$$

Here λ is known as the parameter of the distribution.

We shall use the notation $X \sim P(\lambda)$ to denote that $X$ is a Poisson variate with paramter λ.

**Remarks 1.** It should be noted that

$$\sum_{x=0}^{\infty} P(X = x) = e^{-\lambda} \sum_{x=0}^{\infty} \lambda^x / x! = e^{-\lambda} e^{\lambda} = 1$$

**2.** The corresponding distribution function is:

$$F(x) = P(X \le x) = \sum_{r=0}^{x} p(r) = e^{-\lambda} \sum_{r=0}^{x} \lambda^r / r!\,;\; x = 0, 1, 2, \ldots$$

**3.** Poisson distribution occurs when there are events which do not occur as outcomes of a definite number of trials (unlike that in binomial) of an experiment but which occur at random points of time and space wherein our interest lies only in the number of occurrences of the event, not in its non-occurrences.

**4.** Following are some instances where Poisson distribution may be successfully employed·

(1) Number of deaths from a disease (not in the form of an epidemic) such as heart attack or cancer or due to snake bite.

(2) Number of suicides reported in a particular city.

(3) The number of defective material in a packing manufactured by a good concern.

(4) Number of faulty blades in a packet of 100.

(5) Number of air accidents in some unit of time.

(6) Number of printing mistakes at each page of the book.

(7) Number of telephone calls received at a particular telephone exchange in some unit of time or connections to wrong numbers in a telephone exchange.

(8) Number of cars passing a crossing per minute during the busy hours of a day.

(9) The number of fragments received by a surface area '$t$' from a fragment atom bomb.

(10) The emission of radioactive (alpha) particles.

**7·3·1. The Poisson Process.** The Poisson distribution may also be obtained independently (*i.e.*, without considering it as a limiting form of the Binomial distribution) as follows :

Let $X_t$ be the number of telephone calls received in time interval '$t$' on a telephone switch board. Consider the following experimental conditions :

(1) The probability of getting a call in small time interval $(t, t + dt)$ is $\lambda\, dt$, where $\lambda$ is a positive constant and $dt$ denotes a small increment in time '$t$'.

(2) The probability of getting more than one call in this time interval is very small, *i.e.*, is of the order of $(dt)^2$ *i.e.*, $0\,[(dt)^2]$ such that

$$\lim_{dt \to 0} \frac{0\,(dt)^2}{dt} = 0$$

(3) The probability of any particular call in the time interval $(t, t + dt)$ is independent of the actual time $t$ and also of all previous calls.

Under these conditions it can be shown that the probability of getting $x$ calls in time '$t$', say, $P_x(t)$ is given by

$$P_x(t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}\,;\; x = 0, 1, 2, ..., \infty$$

which is a Poisson distribution with parameter $\lambda t$.

**Proof:** Let $P_x(t) = P$ {of getting $x$ calls in a time interval of length '$t$'}. Also $P$ {of at least one call during $(t, t + dt)$} $= \lambda\, dt + 0\,[(dt)^2]$

and $P$ {of more than one call during $(t, t + dt)$} $= 0\,[(dt)^2]$.

The event of getting exactly $x$ calls in time $t + dt$ can materialise in the following two mutually exclusive ways :

(*i*) $x$ calls in $(0, t)$ and none during $(t, t + dt)$ and the probability of this event is $P_x(t)\{1 - [(\lambda\, dt + 0\,(dt)^2]\}$,

(*ii*) exactly $(x - 1)$ calls during $(0, t)$ and one call in $(t, t + dt)$ and the probability of this event is $P_{x-1}(t)\cdot(\lambda\, dt)$.

Hence by the addition theorem of probability, we get

$$P_x(t + dt) = P_x(t)\{1 - \lambda\, dt - 0\,(dt)^2\} + P_{x-1}(t)\,\lambda\, dt.$$

$$= P_x(t)(1 - \lambda\, dt) + P_{x-1}(t)\,\lambda\, dt + 0\,(dt)^2 P_x(t) \quad ...(1)$$

$$\Rightarrow \frac{P_x(t + dt) - P_x(t)}{dt} = -\lambda P_x(t) + \lambda P_{x-1}(t) + \frac{0\,(dt)^2}{dt} P_x(t)$$

Proceeding to the limit as $dt \to 0$, we get

$$\lim_{dt \to 0} \frac{P_x(t + dt) - P_x(t)}{dt} = -\lambda P_x(t) + \lambda P_{x-1}(t)$$

$$\therefore \quad P_x'(t) = -\lambda P_x(t) + \lambda P_{x-1}(t), x \geq 1 \quad ...(2)$$

where ( $'$ ) denotes differentiation w.r. to '$t$'.

For $x = 0, P_{x-1}(t) = P_{-1}(t) = P\{(-1)$ calls in time 't'$\} = 0$

Hence from (1), we get

$$P_0(t + dt) = P_0(t)\{1 - \lambda\, dt\} + 0\,(dt)^2$$

which on taking the limit $dt \to 0$, gives

$$P_0'(t) = -\lambda P_0(t) \Rightarrow \frac{P_0'(t)}{P_0(t)} = -\lambda$$

Integrating w.r. to. '$t$', we get

$$\log P_0(t) = -\lambda t + C,$$

where $C$ is an arbitrary constant to be determined from the condition

$$P_0(0) = 1$$

Hence $\quad \log 1 = C \Rightarrow C = 0$

$$\therefore \quad \log P_0(t) = -\lambda t \Rightarrow P_0(t) = e^{-\lambda t}$$

Substituting this value of $P_0(t)$ in (2), we get, with $x = 1$

$$P_1'(t) = -\lambda P_1(t) + \lambda e^{-\lambda t}$$

$$\Rightarrow \quad P_1'(t) + \lambda P_1(t) = \lambda e^{-\lambda t}$$

This is an ordinary linear differential equation whose integrating factor is $e^{\lambda t}$. Hence its solution is

$$e^{\lambda t} P_1(t) = \lambda \int e^{\lambda t} e^{-\lambda t}\, dt + C_1 = \lambda t + C_1,$$

where $C_1$ is an arbitrary constant to be determined from $P_1(0) = 0$, which gives $C_1 = 0$.

$$\therefore \qquad P_1(t) = e^{-\lambda t} \lambda t$$

Again substituting this in (2) with $x = 2$, we get

$$P_2'(t) + \lambda P_2(t) = \lambda e^{-\lambda t} \lambda t$$

Integrating factor of this equation is $e^{\lambda t}$ and its solution is

$$P_2(t) e^{\lambda t} = \lambda^2 \int t e^{-\lambda t} e^{\lambda t}\, dt + C_2 = \frac{\lambda^2 t^2}{2} + C_2$$

where $C_2$ is an arbitrary constant to be determined from $P_2(0) = 0$, which gives $C_2 = 0$. Hence

$$P_2(t) = e^{-\lambda t} \frac{(\lambda t)^2}{2}$$

Proceeding similarly step by step, we shall get

$$P_x(t) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}; \; x = 0, 1, 2, \ldots, \infty.$$

### 7·3·2. Moments of the Poisson Distribution

$$\mu_1' = E(X) = \sum_{x=0}^{\infty} x\, p(x, \lambda)$$

$$= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda} \lambda^x}{x!} = \lambda e^{-\lambda} \left[ \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \right]$$

$$= \lambda e^{-\lambda} \left( 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots \right) = \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda$$

Hence the mean of the Poisson distribution is $\lambda$.

$$\mu_2' = E(X^2) = \sum_{x=0}^{\infty} x^2 p(x, \lambda) = \sum_{x=0}^{\infty} \{x(x-1) + x\} \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} x(x-1) \frac{\lambda^x}{x!} + \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= \lambda^2 e^{-\lambda} \left[ \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} \right] + \lambda = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda$$

$$\mu_3' = E(X^3) = \sum_{x=0}^{\infty} x^3 p(x, \lambda)$$

$$= \sum_{x=0}^{\infty} \{x(x-1)(x-2) + 3x(x-1) + x\} \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} x(x-1)(x-2) \frac{e^{-\lambda} \lambda^x}{x!} + 3 \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda} \lambda^x}{x!} + \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= e^{-\lambda}\lambda^3\left[\sum_{x=3}^{\infty}\frac{\lambda^{x-3}}{(x-3)!}\right] + 3e^{-\lambda}\lambda^2\left[\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!}\right] + \lambda$$

$$= e^{-\lambda}\lambda^3 e^{\lambda} + 3e^{-\lambda}\lambda^2 e^{\lambda} + \lambda = \lambda^3 + 3\lambda^2 + \lambda$$

$$\mu_4' = E(X^4) = \sum_{x=0}^{\infty} x^4 \cdot p(x;\lambda)$$

$$= \sum_{x=0}^{\infty}\{x(x-1)(x-2)(x-3) + 6x(x-1)(x-2) + 7x(x-1) + x\}\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\lambda^4\left[\sum_{x=4}^{\infty}\frac{\lambda^{x-4}}{(x-4)!}\right] + 6e^{-\lambda}\lambda^3\left[\sum_{x=3}^{\infty}\frac{\lambda^{x-3}}{(x-3)!}\right]$$

$$+ 7e^{-\lambda}\lambda^2\left[\sum_{x=2}^{\infty}\left(\frac{\lambda^{x-2}}{(x-2)!}\right]\right. + \lambda$$

$$= \lambda^4(e^{-\lambda}e^{+\lambda}) + 6\lambda^3(e^{-\lambda}e^{+\lambda}) + 7\lambda^2(e^{-\lambda}e^{+\lambda}) + \lambda = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$$

The four central moments are now obtained as follows :

$$\mu_2 = \mu_2' - \mu_1'^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$$

Thus the mean and the variance of the Poisson distribution are each equal to $\lambda$.

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^3 = (\lambda^3 + 3\lambda^2 + \lambda) - 3\lambda(\lambda^2 + \lambda) + 2\lambda^3 = \lambda.$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$$

$$= (\lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda) - 4\lambda(\lambda^3 + 3\lambda^2 + \lambda) + 6\lambda^2(\lambda^2 + \lambda) - 3\lambda^4 = 3\lambda^2 + \lambda$$

Co-efficients of skewness and kurtosis are given by

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda} \text{ and } \gamma_1 = \sqrt{\beta_1} = \frac{1}{\sqrt{\lambda}}$$

Also $$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{1}{\lambda} \text{ and } \gamma_2 = \beta_2 - 3 = \frac{1}{\lambda} \quad \text{...(7·15)}$$

Hence the Poisson distribution is always a skewed distribution.

Proceeding to the limit as $\lambda \to \infty$, we get

$$\beta_1 = 0 \text{ and } \beta_2 = 3$$

### 7·3·3. Mode of the Poisson Distribution

$$\frac{p(x)}{p(x-1)} = \frac{\dfrac{e^{-\lambda}\lambda^x}{x!}}{\dfrac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}} = \frac{\lambda}{x} \quad \text{..(7·16)}$$

We discuss the following cases :

**Case I.** *When $\lambda$ is not an integer.*

Let us suppose that S is the intergral part of $\lambda$.

$$\frac{p(1)}{p(0)} > 1, \quad \ldots, \frac{p(S-1)}{p(S-2)} > 1, \frac{p(S)}{p(S-1)} > 1,$$

and $$\frac{p(S+1)}{p(S)} < 1, \frac{p(S+2)}{p(S+1)} < 1, \ldots$$

Combining the above expressions into a single expression, we get

$p(0) < p(1) < p(2) \ldots < p(S-2) < p(S-1) < p(S) > p(S+1) > p(S+2) > \ldots$, which shows that $p(S)$ is the maximum value. Hence in this case the distribution is unimodal and the integral part of $\lambda$ is the unique modal value.

**Case II.** *When $\lambda = k$ (say) is an integer.* Here we have

$$\frac{p(1)}{p(0)} > 1, \frac{p(2)}{p(1)} > 1, \ldots, \frac{p(k-1)}{p(k-2)} > 1$$

and $$\frac{p(k)}{p(k-1)} = 1, \frac{p(k+1)}{p(k)} < 1, \frac{p(k+2)}{p(k+1)} < 1, \ldots$$

$\therefore \; p(0) < p(1) < p(2) < \ldots < p(k-2) < p(k-1) = p(k) > p(k+1) > p(k+2) \ldots$

In this case we have two maximum values, *viz.*, $p(k-1)$ and $p(k)$ and thus the distribution is bimodal and two modes are at $(k-1)$ and $k$, *i.e.*, at $(\lambda - 1)$ and $\lambda$, (since $k = \lambda$).

**7·3·4. Recurrence Relation for the Moments of the Poisson Distribution.** By def.,

$$\mu_r = E\{X - E(X)\}^r = \sum_{x=0}^{\infty} (x-\lambda)^r \, p(x, \lambda)$$

$$= \sum_{x=0}^{\infty} (x-\lambda)^r \frac{e^{-\lambda}\lambda^x}{x!}$$

Differentiating with respect to $\lambda$, we get

$$\frac{d\mu_r}{d\lambda} = \sum_{x=0}^{\infty} r(x-\lambda)^{r-1}(-1)\frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} \frac{(x-\lambda)^r}{x!}\left\{x\lambda^{x-1}e^{-\lambda} - \lambda^x e^{-\lambda}\right\}$$

$$= -r\sum_{x=0}^{\infty} (x-\lambda)^{r-1} \cdot \frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} \frac{(x-\lambda)^r}{x!}\left\{\lambda^{x-1}e^{-\lambda}(x-\lambda)\right\}$$

$$= -r\sum_{x=0}^{\infty} (x-\lambda)^{r-1}\frac{e^{-\lambda}\lambda^x}{x!} + \frac{1}{\lambda}\sum_{x=0}^{\infty} (x-\lambda)^{r+1} \cdot \frac{e^{-\lambda}\lambda^x}{x!}$$

$$\therefore \quad \frac{d\mu_r}{d\lambda} = -r\mu_{r-1} + \frac{1}{\lambda}\mu_{r+1}$$

$$\Rightarrow \quad \mu_{r+1} = r\lambda\mu_{r-1} + \lambda\frac{d\mu_r}{d\lambda} \qquad \ldots(7\cdot17)$$

Putting $r = 1, 2$ and $3$ successively, we get

$$\mu_2 = r\mu_o + \lambda\frac{d\mu_1}{d\lambda} = \lambda \qquad (\because \mu_o = 1, \mu_1 = 0)$$

$$\mu_3 = 2\lambda\mu_1 + \lambda\frac{d\mu_2}{d\lambda} = \lambda,\ \mu_4 = 3\lambda\mu_2 + \lambda\frac{d\mu_3}{d\lambda} = 3\lambda^2 + \lambda$$

**7·3·5. Moment Generating Function of the Poisson Distribution**

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx}\cdot\frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty}\frac{e^{-\lambda}(\lambda e^t)^x}{x!}$$

$$= e^{-\lambda}\left\{1 + \lambda e^t + \frac{(\lambda e^t)^2}{2!} + \dots\right\} = e^{-\lambda}\cdot e^{\lambda e^t} = e^{\lambda(e^t-1)} \qquad ...(7\cdot 18)$$

**7·3·6. Characteristic Function of the Poisson Distribution**

$$\phi_X(t) = \sum_{x=0}^{\infty} e^{itx}\cdot p(x) = \sum_{x=0}^{\infty} e^{itx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^{it})^x}{x!}$$

$$= e^{-\lambda}\, e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)} \qquad ...(7\cdot 19)$$

**7·3·7. Cumulants of the Poisson Distribution**

$$K_X(t) = \log M_X(t) = \log\left[e^{\lambda(e^t-1)}\right] = \lambda(e^t - 1)$$

$$= \lambda\left[\left(1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \dots + \frac{t^r}{r!} + \dots\right) - 1\right]$$

$$= \lambda\left[t + \frac{t^2}{2!} + \frac{t^3}{3!} + \dots + \frac{t^r}{r!} + \dots\right]$$

$\kappa_r$ = $r$th cumulant = co-efficient of $\frac{t^r}{r!}$ in $K_X(t) = \lambda$

$$\Rightarrow \qquad \kappa_r = \lambda;\ r = 1, 2, 3, \dots \qquad ...(7\cdot 19a)$$

Hence all the cumulants of the Poisson distribution are equal, each being equal to $\lambda$. In particular, we have

$$\text{Mean} = \kappa_1 = \lambda,\ \mu_2 = \kappa_2 = \lambda,\ \mu_3 = \kappa_3 = \lambda \text{ and } \mu_4 = \kappa_4 + 3\kappa_2^2 = \lambda + 3\lambda^2$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda} \text{ and } \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\lambda + 3\lambda^2}{\lambda^2} = \frac{1}{\lambda} + 3$$

**Remark.** If $m$ is the mean and $\sigma$ is the s.d. of Poisson distribution with parameter $\lambda$, then

$$m\,\sigma\,\gamma_1\,\gamma_2 = \lambda\cdot\sqrt{\lambda}\cdot\sqrt{\beta_1}\,(\beta_2 - 3)$$

$$= \lambda\cdot\sqrt{\lambda}\cdot\frac{1}{\sqrt{\lambda}}\cdot\frac{1}{\lambda} = 1.$$

**7·3·8. Additive or Reproductive Property of Independent Poisson Variates.** *Sum of independent Poisson variates is also a Poisson variate. More elaborately, if $X_i$, $(i = 1, 2, \dots, n)$ are independent Poisson variates with param-*

*ters* $\lambda_i$ ; $i = 1, 2, ..., n$ *respectively, then* $\sum_{i=1}^{n} X_i$ *is also a Poisson variate with parameter* $\sum_{i=1}^{n} \lambda_i$.

**Proof.** $M_{X_i}(t) = e^{\lambda_i(e^t - 1)}$ ; $i = 1, 2, ..., n$

$$M_{X_1 + X_2 + \ldots + X_n}(t) = M_{X_1}(t)\, M_{X_2}(t) \ldots M_{X_n}(t),$$

[since $X_i$ ; $i = 1, 2, ..., n$ are independent]

$$= e^{\lambda_1(e^t - 1)}\, e^{\lambda_2(e^t - 1)} \ldots e^{\lambda_n(e^t - 1)}$$

$$= e^{(\lambda_1 + \lambda_2 + \ldots + \lambda_n)(e^t - 1)}$$

which is the m.g.f. of a Poisson variate with parameter $\lambda_1 + \lambda_2 + \ldots + \lambda_n$. Hence by uniqueness theorem of m.g.f.'s, $\sum_{i=1}^{n} X_i$ is also a Poisson variate with parameter $\sum_{i=1}^{n} \lambda_i$.

**Remarks 1. In fact, the converse of the above result is also true** *i.e., If* $X_1, X_2, ..., X_n$ *are independent and* $\sum_{i=1}^{n} X_i$ *has a Poisson distribution, then each of the random variables* $X_1, X_2, ..., X_n$ *has a Poisson distribution.*

Let $X_1$ and $X_2$ be independent *r.v.'s* so that $X_1 \sim P(\lambda_1)$ and $X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$. Then we want to prove that $X_2 \sim P(\lambda_2)$.

**Proof.** Since $X_1$ and $X_2$ are independent, we have

$$M_{X_1 + X_2}(t) = M_{X_1}(t)\, M_{X_2}(t)$$

$$\Rightarrow \quad e^{(\lambda_1 + \lambda_2)(e^t - 1)} = e^{\lambda_1(e^t - 1)} \cdot M_{X_2}(t)$$

$$\Rightarrow \quad M_{X_2}(t) = e^{\lambda_2(e^t - 1)}$$

$\Rightarrow$ $X_2 \sim P(\lambda_2)$, by uniqueness theorem of m.g.f.

**2.** The difference of two independent Poisson variates is not a Poisson variate.

$$M_{X_1 - X_2}(t) = M_{X_1 + (-X_2)}(t) = M_{X_1}(t) \cdot M_{(-X_2)}(t),$$

(since $X_1$ and $X_2$ are independent).

$$\therefore \quad M_{X_1 - X_2}(t) = M_{X_1}(t)\, M_{X_2}(-t) \qquad [\because M_{cX}(t) = M_X(ct)]$$

$$= e^{\lambda_1(e^t - 1)} \cdot e^{\lambda_2(e^{-t} - 1)} = e^{\lambda_1(e^t - 1) + \lambda_2(e^{-t} - 1)}$$

which cannot be put in the form $e^{\lambda(e^t - 1)}$ Hence $(X_1 - X_2)$ is not a Poisson variate.

Moreover the difference $(X_1 - X_2)$ cannot be a Poisson variate is evident from the fact that it may have positive as well as negative values, while a Poisson variate is always non-negative.

**7·3·9. Probability Generating Function of Poisson Distribution**

$$\text{P.G.F. of } X = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}\cdot s^k = \sum_{k=0}^{\infty} e^{-\lambda}\frac{(\lambda s)^k}{k!} = e^{-\lambda}e^{\lambda s} = e^{\lambda(s-1)} \qquad ...(7\cdot 20)$$

**Example 7·24.** *A car hire firm has two cars which it fires out day by day. The number of demands for a car on each day is distributed as Poisson variate with mean 1·5. Calculate the proportion of days on which (i) neither car is used, and (ii) some demand is refused.* **[Meerut Univ. B.Sc. 1993]**

**Solution.** The proportion of days on which there are $x$ demands for a car

$$= P\{\text{of } x \text{ demands in a day}\} = \frac{e^{-1\cdot 5}(1\cdot 5)^x}{x!},$$

since the number of demands for a car on any day is a Poisson variate with mean 1·5. Thus

$$P(X=x) = \frac{e^{-1\cdot 5}(1\cdot 5)^x}{x!};\quad x = 0, 1, 2, \ldots$$

(*i*) Proportion of days on which neither car is used is given by

$$P(X=0) = e^{-1\cdot 5}$$
$$= \left[1 - 1\cdot 5 + \frac{(1\cdot 5)^2}{2!} - \frac{(1\cdot 5)^3}{3!} + \frac{(1\cdot 5)^4}{4!} - \ldots\right]$$
$$= 0\cdot 2231$$

(*ii*) Proportion of days on which some demand is refused is

$$P(X>2) = 1 - P(X \le 2)$$
$$= 1 - [P(X=0) + P(X=1) + P(X=2)]$$
$$= 1 - e^{-1\cdot 5}\left[1 + 1\cdot 5 + \frac{(1\cdot 5)^2}{2!}\right]$$
$$= 1 - 0\cdot 2231 \times 3\cdot 625 = 0\cdot 19126$$

**Example 7·25.** *A manufacturer of cotter pins knows that 5% of his product is defective. If he sells cotter pins in boxes of 100 and guarantees that not more than 10 pins will be defective, what is the approximate probability that a box will fail to meet the guaranteed quality ?* **[Kanpur Univ. B.Sc. 1993]**

**Solution.** We are given $n = 100$.

Let $p$ = Probability of a defective pin = 5% = 0·05

$\therefore$ $\lambda$ = Mean number of defective pins in a box of 100

$= np = 100 \times 0\cdot 05 = 5$

Since '$p$' is small, we may use Poisson distribution.

Probability of $x$ defective pins in a box of 100 is

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-5}5^x}{x!};\ x = 0, 1, 2, \ldots$$

Probability that a box will fail to meet the guaranteed quality is

$$P(X > 10) = 1 - P(X \le 10) = 1 - \sum_{x=0}^{10} \frac{e^{-5}5^x}{x!} = 1 - e^{-5} \sum_{x=0}^{10} \frac{5^x}{x!}$$

**Example 7·26.** *Six coins are tossed 6,400 times. Using the Poisson distribution, find the approximate probability of getting six heads r times.*

**Solution.** The probability of obtaining six heads in one throw of six coins (*a* single trial), is $p = (½)^6$, assuming that head and tail are equally probable.

$$\therefore \qquad \lambda = np = 6400 \times (½)^6 = 100.$$

Hence, using Poisson probability law, the required probability of getting 6 heads $r$ times is given by :

$$P(X = r) = \frac{e^{-\lambda} \cdot \lambda^r}{r!} = \frac{e^{-100} \cdot (100)^r}{r!};\ r = 0, 1, 2, \ldots$$

**Example 7·27.** *In a book of 520 pages, 390 typo-graphical errors occur. Assuming Poisson law for the number of errors per page, find the probability that a random sample of 5 pages will contain no error.*

**[Patna Univ. B.Sc. (Hons.), 1988]**

**Solution.** The average number of typographical errors per page in the book is given by $\lambda = (390/520) = 0 \cdot 75$

Hence using Poisson probability law, the probability of $x$ errors per page is given by : $P(X = x) = \dfrac{e^{-\lambda}\lambda^x}{x!} = \dfrac{e^{-0\cdot 75}(0\cdot 75)^x}{x!};\ x = 0, 1, 2, \ldots$

The required probability that a random sample of 5 pages will contain no error is given by : $[P(X = 0)]^5 = \left(e^{-0\cdot 75}\right)^5 = e^{-3\cdot 75}$

**Example 7·28.** *Suppose that the number of telephone calls coming into a telephone exchange between 10 A.M. and 11 A.M. say, $X_1$ is a random variable with Poisson distribution with parameter 2. Similarly the number of calls arriving between 11 A.M. and 12 noon say, $X_2$ has a Poisson distribution with parameter 6. If $X_1$ and $X_2$ are independent, what is the probability that more than 5 calls come in between 10 A.M. and 12 noon ?* **[Calicut U. B. Sc. Oct. 1992]**

**Solution.** Let $X = X_1 + X_2$. By the additive property of Poisson distribution, $X$ is also a Poisson variate with parameter (say) $\lambda = 2 + 6 = 8$

Hence the probability of $x$ calls in-between 10. A.M. and 12 noon is given by $P(X = x) = \dfrac{e^{-\lambda}\lambda^x}{x!} = \dfrac{e^{-8}8^x}{x!};\ x = 0, 1, 2, \ldots$

Probability that more than 5 calls come in between 10 A.M. and 12 noon is given by

$$P(X > 5) = 1 - P(X \le 5) = 1 - \sum_{x=0}^{5} \frac{e^{-8} 8^x}{x!}$$
$$= 1 - 0 \cdot 1912 = 0 \cdot 8088$$

**Example 7·29.** *A Poisson distribution has a double mode at $x = 1$ and $x = 2$. What is the probability that x will have one or the other of these two values?*

**Solution.** We have proved that if the Poisson distribution is bimodal, then the two modes are at the points $x = \lambda - 1$ and $x = \lambda$. Since we are given that the two modes are at the points $x = 1$ and $x = 2$, we find that $\lambda = 2$.

$$\therefore \qquad P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} = \frac{e^{-2} 2^x}{x!}; \; x = 0, 1, 2, \ldots$$
$$\Rightarrow \qquad P(X = 1) = e^{-2} 2$$
and
$$P(X = 2) = \frac{e^{-2} \cdot 2^2}{2!} = e^{-2} \cdot 2$$

Required probability $= P(X = 1) + P(X = 2) = 2e^{-2} + 2e^{-2} = 0{\cdot}542$

**Example 7·30.** *If $X$ is a Poisson variate such that*

$$P(X = 2) = 9\,P(X = 4) + 90\,P(X = 6) \qquad \ldots(*)$$

*Find (i) $\lambda$, the mean of $X$, (ii) $\beta$, the coefficient of skewness.*

**[Delhi Univ. B. Sc. (Maths. Hons.) 1992, '87]**

**Solution.** If $X$ is a Poisson variate with parameter $\lambda$, then

$$P(X = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, \; x = 0, 1, 2, \ldots; \lambda > 0$$

Hence (*) gives

$$\frac{e^{-\lambda} \cdot \lambda^2}{2!} = e^{-\lambda} \left[ 9 \, \frac{\lambda^4}{4!} + 90 \, \frac{\lambda^6}{6!} \right]$$
$$= \frac{e^{-\lambda} \lambda^2}{8} \left[ 3\lambda^2 + \lambda^4 \right]$$
$$\Rightarrow \qquad \lambda^4 + 3\lambda^2 - 4 = 0$$

Solving as a quadratic in $\lambda^2$, we get

$$\lambda^2 = \frac{-3 \pm \sqrt{9 + 16}}{2} = \frac{-3 \pm 5}{2}$$

Since $\lambda > 0$, we get $\lambda^2 = 1 \quad \Rightarrow \quad \lambda = 1$

Hence mean $= \lambda = 1$, and $\mu_2 =$ Variance $= \lambda = 1$

Also $\quad \beta_1 =$ Coefficient of skewness $= \dfrac{1}{\lambda} = 1.$

**Example 7·31.** *If $X$ and $Y$ are independent Poisson variates such that*

$$P(X = 1) = P(X = 2)$$
*and*
$$P(Y = 2) = P(Y = 3) \qquad \ldots(*)$$

*Find the varaince of $X - 2Y$.*

**Solution.** Let $X \sim P(\lambda)$ and $Y \sim P(\mu)$.

Then we have

$$P(X = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, \; x = 0, 1, 2, \ldots ; \lambda > 0$$

and $$P(Y = y) = \frac{e^{-\mu} \cdot \mu^y}{y!}, \; y = 0, 1, 2, \ldots ; \mu > 0$$

Using (*), we get

$$\left.\begin{aligned} \lambda e^{-\lambda} &= \frac{\lambda^2 e^{-\lambda}}{2!} \\ \text{and} \quad \frac{\mu^2 e^{-\mu}}{2} &= \frac{\mu^3 e^{-\mu}}{3!} \end{aligned}\right\} \quad \ldots(**)$$

Solving (**), we get

$$\lambda e^{-\lambda} [\lambda - 2] = 0 \text{ and } \mu^2 e^{-\mu} [\mu - 3] = 0$$

$\Rightarrow$ $\lambda = 2$ and $\mu = 3$, since $\lambda > 0$, $\mu > 0$.

Now $\quad \text{Var}(X) = \lambda = 2$, and $\text{Var}(Y) = \mu = 3 \quad \ldots(***)$

$\therefore \quad \text{Var}(X - 2Y) = 1^2 \, \text{Var}(X) + (-2)^2 \cdot \text{Var}\, Y,$

covariance term vanishes since $X$ and $Y$ are independent.

Hence, on using (***), we get

$$\text{Var}(X - 2Y) = 2 + 4 \times 3 = 14$$

**Example 7·32.** *If $X$ and $Y$ are independent Poisson variates with means $\lambda_1$ and $\lambda_2$ respectively, find the probability that*

(*i*) $X + Y = k$, (*ii*) $X = Y$ [Delhi Univ. B. Sc. (Stat. Hons.), 1991]

**Solution.** We have

$$P(X = x) = \frac{e^{-\lambda_1} \cdot \lambda_1^x}{x!}, \; x = 0, 1, 2, 3, \ldots ; \lambda_1 > 0$$

and $$P(Y = y) = \frac{e^{-\lambda_2} \cdot \lambda_2^y}{y!}, \; y = 0, 1, 2, 3 \ldots; \lambda_2 > 0$$

(*i*) $$P(X + Y = k) = \sum_{r=0}^{k} P(X = r \cap Y = k - r)$$

$$= \sum_{r=0}^{k} P(X = r) \, P(Y = k - r)$$

[$\because$ $X$ and $Y$ are independent]

$$= \sum_{r=0}^{k} \frac{e^{-\lambda_1} \lambda_1^r}{r!} \cdot \frac{e^{-\lambda_2} \cdot \lambda_2^{k-r}}{(k-r)!}$$

$$= e^{-(\lambda_1 + \lambda_2)} \sum_{r=0}^{k} \frac{\lambda_1^r \cdot \lambda_2^{k-r}}{r!\,(k-r)!}$$

$$= e^{-(\lambda_1+\lambda_2)}\left[\frac{\lambda_2^k}{k!}+\frac{\lambda_1\cdot\lambda_2^{k-1}}{1!\,(k-1)!}+\frac{\lambda_1^2\cdot\lambda_2^{k-2}}{2!\,(k-2)!}+\ldots+\frac{\lambda_1^k}{k!}\right]$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{k!}\left[\lambda_2^k + {}^kC_1\,\lambda_2^{k-1}\cdot\lambda_1 + {}^kC_2\cdot\lambda_2^{k-2}\cdot\lambda_1^2+\ldots+\lambda_1^k\right]$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{k!}\times(\lambda_1+\lambda_2)^k,\ k = 0, 1, 2, \ldots$$

which is the probability function of Poisson distribution with parameter $\lambda_1 + \lambda_2$.

**Aliter.** Since $X \sim P(\lambda_1)$ and $Y \sim P(\lambda_2)$ are independent, by the additive property of Poisson distribution $X + Y \sim P(\lambda_1 + \lambda_2)$. Hence

$$P(X+Y=k) = \frac{e^{-(\lambda_1+\lambda_2)}\times(\lambda_1+\lambda_2)^k}{k!};\ k = 0, 1, 2, \ldots$$

*(ii)* $$P(X=Y) = \sum_{r=0}^{\infty} P(X=r \cap Y=r)$$

$$= \sum_{r=0}^{\infty} P(X=r)\,P(Y=r)$$

[$\because X$ and $Y$ are independent]

$$= e^{-(\lambda_1+\lambda_2)}\sum_{r=0}^{\infty}\frac{(\lambda_1\lambda_2)^r}{(r!)^2}$$

**Example 7·33.** *Show that in a Poisson distribution with unit mean, mean deviation about mean is (2/e) times the standard deviation.*

**[Patna Univ. B. Sc. (Stat. Hons.) 1992; Delhi Univ. B.Sc. (Stat. Hons.), 1993]**

**Solution.** Here we are given $\lambda = 1$.

$$\therefore \quad P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-1}\cdot 1}{x!} = \frac{e^{-1}}{x!};\ x = 0, 1, 2, \ldots$$

Mean deviation about mean 1 is

$$E(|X-1|) = \sum_{x=0}^{\infty}|x-1|\,p(x) = e^{-1}\sum_{x=0}^{\infty}\frac{|x-1|}{x!}$$

$$= e^{-1}\left[1+\frac{1}{2!}+\frac{2}{3!}+\frac{3}{4!}+\ldots\right]$$

We have $$\frac{n}{(n+1)!} = \frac{(n+1)-1}{(n+1)!} = \frac{1}{n!}-\frac{1}{(n+1)!}$$

$\therefore$ Mean deviation about mean

$$= e^{-1}\left[1 + \left(1 - \frac{1}{2!}\right) + \left(\frac{1}{2!} - \frac{1}{3!}\right) + \left(\frac{1}{3!} - \frac{1}{4!}\right) + \ldots\right]$$

$$= e^{-1}(1 + 1) = \frac{2}{e} \times 1 = \frac{2}{e} \times \text{ standard deviation,}$$

since for the Poisson distribution, variance = mean = 1 (given).

**Example 7·34.** *Let $X_1, X_2, \ldots, X_n$ be identically and independently distributed Bin (1, p) variates. Let $S_n = \sum_{j=1}^{n} X_j$ and $M_n(t)$ be the m.g.f. of $S_n$. Find $\lim_{n \to \infty} M_n(t)$, using $np = \lambda$ (const.)* **[Delhi Univ. B. Sc. (Maths Hons.), 1989]**

**Solution.** Since $X_i$, $i = 1, 2, \ldots n$ are i.i.d. binomial variates $B(1, p)$,

$S_n = \sum_{j=1}^{n} X_j$, is a binomial $B(n, p)$ variate.

$$\therefore\ M_n(t) = M.g.f. \text{ of } S_n = \left(q + pe^t\right)^n = \left[1 + (e^t - 1)p\right]^n$$

If we take $np = \lambda \Rightarrow p = \lambda/n$ and let $n \to \infty$, we get

$$\lim_{n \to \infty} M_n(t) = \lim_{n \to \infty}\left[1 + \frac{(e^t - 1)\lambda}{n}\right]^n = \exp[\lambda(e^t - 1)],$$

which is the m.g.f. of Poisson distribution with parameter $\lambda$. Hence by uniqueness theorem of m.g.f., $S_n = \sum_{j=1}^{n} X_j \to P(\lambda)$, as $n \to \infty$, with $np = \lambda$ (fixed).

**Example. 7·35.** *(a) If $X$ is a Poisson variate with mean $m$, show that the expectation of $e^{-kX}$ is $\exp[-m(1 - e^{-k})]$.* **[Nagpur Univ. B.Sc. 1993]**

*Hence show that, if $\bar{X}$ is the arithmetic mean of $n$ independent random variables $X_1, X_2, \ldots, X_n$, each having Poisson distribution with parameter $m$, then $e^{-\bar{x}}$ as an estimate of $e^{-m}$ is biased, although $\bar{X}$ is an unbaised estimate of $m$.*

*(b) If $X$ is a Poisson variate with mean $m$, what would be the expectation of $e^{-kx} kX$, $k$ being a constant.*

**Solution.**

$$E(e^{-kX}) = \sum_{x=0}^{\infty} e^{-kx} p(x) = \sum_{x=0}^{\infty} e^{-kx} \cdot \frac{e^{-m} m^x}{x!} = e^{-m} \sum_{x=0}^{\infty} \frac{(me^{-k})^x}{x!}$$

$$= e^{-m}\left[1 + me^{-k} + \frac{(me^{-k})^2}{2!} + \ldots\right]$$

$$= e^{-m}\, e^{me^{-k}} = e^{-m(1-e^{-k})} \qquad \text{...(*)}$$

We have

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$$

Since $X_i$; $i = 1, 2, ..., n$ is a Poisson variate with parameter $m$, $E(X_i) = m$.

$$\therefore\ E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} m = \frac{1}{n}\, nm = m$$

Hence $\bar{X}$ is an unbaised estimate of $m$.

Now $$E(e^{-\bar{X}}) = E\left[\exp\left(\frac{-1}{n}\sum_{i=1}^{n} X_i\right)\right]$$
$$= E\left(e^{-X_1/n} . e^{-X_2/n} \ldots e^{-X_n/n}\right),$$
$$= E(e^{-X_1/n})\, E(e^{-X_2/n}) \ldots E(e^{-X_n/n}),$$

(since $X_1, X_2, ..., X_n$ are independent)

$$\therefore\quad E(e^{-\bar{X}}) = \prod_{i=1}^{n} E(e^{-X_i/n}) \qquad \text{...(**)}$$

Using (*) with $k = 1/n$, we get

$E(e^{-X_i/n}) = e^{-m(1-e^{-1/n})}$, (since $X_i$ is a Poisson variate with parameter $m$)

$$\therefore\ E(e^{-\bar{X}}) = \prod_{i=1}^{n}\left[\exp\left\{-m(1-e^{-1/n})\right\}\right] = \left[\exp\left\{-m(1-e^{-1/n})\right\}\right]^{n}$$
$$= \exp\left\{-mn(1-e^{-1/n})\right\} \neq e^{-m}$$

Hence $e^{-X}$ is not an unbaised estimated of $e^{-m}$, though $\bar{X}$ is an unbiased estimate of $m$.

$$(b)\ E(e^{-kX} kX) = \sum_{x=0}^{\infty} e^{-kx} kx \cdot p(x) = k\sum_{x=1}^{\infty} e^{-kx}\, x\, \frac{e^{-m} m^x}{x!}$$
$$= ke^{-m}\sum_{x=1}^{\infty}\frac{(me^{-k})^x}{(x-1)!} = ke^{-m} me^{-k}\sum_{x=1}^{\infty}\frac{(me^{-k})^{x-1}}{(x-1)!}$$
$$= mke^{-m-k}\left\{1 + me^{-k} + \frac{(me^{-k})^2}{2!} + \ldots\right\}$$
$$= mk\,e^{-m-k}\cdot e^{me^{-k}} = mk\exp\left[\left\{m(e^{-k}-1)\right\} - k\right]$$

**Example 7·36.** *If $X$ and $Y$ are independent Poisson variates with means $m_1$ and $m_2$ respectively, prove that the probability that $X - Y$ has the value 'r' is the co-efficient of $t^r$ in*

$$\exp\left\{m_1 t + m_2 t^{-1} - m_1 - m_2\right\}$$

**[Delhi Univ. B.Sc. (Stat. Hons.), 1991, '89]**

**Solution.** Since $X$ and $Y$ are independent Poisson variates with means $m_1$ and $m_2$ respectively,

$$\left.\begin{array}{ll} & P(X = x) = \dfrac{e^{-m_1} m_1^x}{x!}\,;\ x = 0, 1, 2, \ldots \infty \\ \text{and} & P(Y = y) = \dfrac{e^{-m_2} m_2^y}{y!}\,;\ y = 0, 1, 2, \ldots \infty \end{array}\right\} \quad \ldots(1)$$

$$P(X - Y = r) = \sum_{s=0}^{\infty} P(X = r + s \cap Y = s) = \sum_{s=0}^{\infty} P(X = r + s)\, P(Y = s)$$

$$= \sum_{s=0}^{\infty} \frac{e^{-m_1} \cdot m_1^{r+s}}{(r+s)!} \cdot \frac{e^{-m_2} m_2^s}{s!} \quad \ldots[\text{From } (1)]$$

$$= e^{-m_1 - m_2} \sum_{s=0}^{\infty} \frac{m_1^{r+s} m_2^s}{(r+s)!\, s!} \quad \ldots(2)$$

We have $e^{m_1 t + m_2 t^{-1}} = e^{m_1 t} \times e^{m_2 t^{-1}}$

$$= \left\{1 + m_1 t + \frac{(m_1 t)^2}{2!} + \ldots + \frac{(m_1 t)^{r+s}}{(r+s)!} + \ldots\right\}$$

$$\times \left\{1 + m_2 t^{-1} + \frac{(m_2 t^{-1})^2}{2!} + \ldots + \frac{(m_2 t^{-1})^s}{s!} + \ldots\right\}$$

$\therefore$ Co-efficient of $t^r$ in $e^{m_1 t + m_2 t^{-1}} = \displaystyle\sum_{s=0}^{\infty} \frac{m_1^{r+s} m_2^s}{(r+s)!\, s!}$

Hence from (2), we get

$$P(X - Y = r) = e^{-m_1 - m_2} \times \text{Coefficient of } t^r \text{ in } e^{m_1 t + m_2 t^{-1}},$$

$$= \text{Coefficient of } t^r \text{ in } e^{-m_1 - m_2 + m_1 t + m_2 t^{-1}}$$

which is the required result.

**Example 7·37.** *If $X$ is a Poisson variate with mean $m$, show that $\dfrac{X - m}{\sqrt{m}}$ is a variable with mean zero and variance unity. Find the M.G.F. for this variable and show that it approaches $e^{t^2/2}$ as $m \to \infty$. Also interpret the result.*

**[Delhi Univ. B. Sc. (Stat. Hons.), 1987]**

**Solution.** Let $Y = \dfrac{X - m}{\sqrt{m}}$

$$\therefore \quad E(Y) = E\left(\frac{X - m}{\sqrt{m}}\right) = \frac{1}{\sqrt{m}} E(X - m) = 0$$

$$V(Y) = E\left(\frac{X-m}{\sqrt{m}}\right)^2 = \frac{1}{m}E(X-m)^2 = \frac{1}{m}\mu_2 = 1$$

$$
\begin{aligned}
M.G.F. \text{ of } Y = M_Y(t) &= E\left(e^{tY}\right) = E\left[e^{t(X-m)/\sqrt{m}}\right] \\
&= e^{-t\sqrt{m}}\left[E\left(e^{tX/\sqrt{m}}\right)\right] \\
&= e^{-t\sqrt{m}} \sum_{x=0}^{\infty} \frac{e^{-m} m^x}{x!} \cdot e^{tx/\sqrt{m}} \\
&= e^{-t\sqrt{m}} \cdot e^{-m} \sum_{x=0}^{\infty} \frac{(me^{t/\sqrt{m}})^x}{x!} \\
&= e^{-m-t\sqrt{m}}\left[1 + \frac{me^{t/\sqrt{m}}}{1!} + \frac{(me^{t/\sqrt{m}})^2}{2!} + \dots\right] \\
&= e^{-m-t\sqrt{m}} \cdot \exp\left(me^{t/\sqrt{m}}\right) = \exp\left[-m - t\sqrt{m} + me^{t/\sqrt{m}}\right] \\
&= \exp\left[-m - t\sqrt{m} + m\left(1 + \frac{t}{\sqrt{m}} + \frac{t^2}{2!\,m} + \frac{t^3}{3!\,m^{3/2}} + \dots\right)\right] \\
&= \exp\left[\frac{1}{2}t^2 + \frac{1}{3!}\frac{t^3}{\sqrt{m}} + \dots\right]
\end{aligned}
$$

Now proceeding to limit as $m \to \infty$, we get

$$\lim_{m \to \infty} M_Y(t) = e^{t^2/2} \qquad \dots(*)$$

**Interpretation.** (*) is the m.g.f. of Standard Normal Variate [c.f. Remark to § 8·2·5]. Hence by uniqueness theorem of m.g.f.'s, standard Poisson variate tends to standard normal variate as $m \to \infty$. Hence Poisson distribution tends to Normal distribution for large values of parameter $m$.

**Example 7·38.** *Deduce the first four moments about the mean of the Poisson distribution from those of the Binomial distribution.*

**Solution.** The first four central moments of the binomial distribution are

$$\left.\begin{array}{ll} \mu_1 = 0, & \text{Mean} = np \\ \mu_2 = npq, & \mu_3 = npq(q-p) \text{ and} \\ \mu_4 = npq(1-6pq) + 3n^2p^2q^2 & \end{array}\right\} \qquad \dots(*)$$

Poisson distribution is a limiting form of the binomial distribution under the following conditions :

(i) $n \to \infty$, (ii) $p \to 0$, *i.e.*, $q \to 1$, and (iii) $np = \lambda$, (say), is finite.

Using these conditions, we get from (*) the moments of the Poisson distribution as

$$
\begin{aligned}
\mu_1 &= 0 \\
\text{Mean} &= \lim(np) = \lambda \\
\mu_2 &= \lim(npq) = \lim(np) \cdot \lim(q) = \lambda \cdot 1 = \lambda \\
\mu_3 &= \lim[npq(q-p)] = \lambda \cdot 1(1-0) = \lambda \\
\mu_4 &= \lim[npq(1-6pq) + 3(np)^2q^2]
\end{aligned}
$$

$$= [\lambda \cdot 1\,(1 - 6\cdot 0\cdot 1) + 3\lambda^2 \cdot 1] = \lambda + 3\lambda^2$$

**Example 7·39.** *If X is a Poisson variate with parameter m and Y is another discrete variable whose conditional distribution for a given X is given by*

$$P(Y = r \mid X = x) = \binom{x}{r} p^r (1-p)^{x-r}\,;\ 0 < p < 1,\ r = 0, 1, 2, \ldots, x$$

*then show that the unconditional distribution of Y is a Poisson distribution with parameter mp.*

**[Delhi Univ. B.Sc. (Stat. Hons.), 1993, Shivaji U.B.Sc. Nov. 1992]**

**Solution.** We are given that

$$P(X = x) = \frac{e^{-m} m^x}{x\,!}\,;\ x = 0, 1, 2, \ldots$$

and
$$P(Y = r \mid X = x) = \binom{x}{r} p^r (1-p)^{x-r}\,;\ r \le x$$

$$\therefore\quad P(X = x \cap Y = r) = P(X = x)\,P(Y = r \mid X = x)$$

$$= \frac{e^{-m} m^x}{x\,!}\binom{x}{r} p^r (1-p)^{x-r}$$

$\therefore\ P(Y = r) =$ The unconditional distribution of $Y$.

$$= \sum_{x=r}^{\infty}\left[\frac{e^{-m} m^x}{x\,!}\cdot\binom{x}{r} p^r (1-p)^{x-r}\right]$$

$$= e^{-m}\left[\sum_{x=r}^{\infty}\binom{x}{r}\frac{p^r m^x (1-p)^{x-r}}{x\,!}\right]$$

$$= e^{-m}\left[\sum_{x=r}^{\infty}\frac{m^x}{x\,!}\cdot\frac{x\,!}{r\,!\,(x-r)\,!}\,p^r (1-p)^{x-r}\right]$$

$$= \frac{e^{-m}}{r\,!}\left[\sum_{x=r}^{\infty}\frac{m^x}{(x-r)\,!}\,p^r (1-p)^{x-r}\right]$$

$$= \frac{e^{-m}(mp)^r}{r\,!}\left[\sum_{x=r}^{\infty}\frac{m^{x-r}(1-p)^{x-r}}{(x-r)\,!}\right]$$

$$= \frac{e^{-m}(mp)^r}{r\,!}\left[\sum_{x=r}^{\infty}\frac{\{m(1-p)\}^{x-r}}{(x-r)\,!}\right]$$

$$= \frac{e^{-m}(mp)^r}{r\,!}\,e^{m(1-p)} = \frac{e^{-mp}(mp)^r}{r\,!}\,;\ r = 0, 1, 2, \ldots$$

Hence $Y$ is a Poisson variate with parameter $mp$.

**Example 7·40.** *If $X$ and $Y$ are independent Poisson variates, show that the conditional distribution of $X$ given $X + Y$, is binomial.*

**[Madras Univ. B.Sc. Main 1992; Delhi Univ. B. Sc. (Maths Hons.), 1988]**

**Solution.** Let $X$ and $Y$ be independent Poisson variates with parameters $\lambda$ and $\mu$ respectively. Then $X + Y$ is also a Poisson variate with parameter $\lambda + \mu$.

$$P[X = r \mid (X + Y = n)] = \frac{P(X = r \cap X + Y = n)}{P(X + Y = n)} = \frac{P(X = r \cap Y = n - r)}{P(X + Y = n)}$$

$$= \frac{P(X = r)P(Y = n - r)}{P(X + Y = n)} \qquad \text{[since } X \text{ and } Y \text{ are indepdent]}$$

$$\therefore P[X = r \mid (X + Y = n)] = \frac{e^{-\lambda}\dfrac{\lambda^r}{r!} \cdot e^{-\mu}\dfrac{\mu^{n-r}}{(n-r)!}}{\dfrac{e^{-(\lambda+\mu)}(\lambda+\mu)^n}{n!}}$$

$$= \frac{n!}{r!(n-r)!}\left(\frac{\lambda}{\lambda+\mu}\right)^r\left(\frac{\mu}{\lambda+\mu}\right)^{n-r}$$

$$= \binom{n}{r} p^r q^{n-r}, \text{ where } p = \frac{\lambda}{\lambda+\mu}, q = 1 - p$$

Hence the conditional distribution of $X$ given $X + Y = n$, is a binomial distribution with parameters $n$ and $p = \lambda/(\lambda + \mu)$.

**Example 7·41.** *If $X$ is a Poisson variate with parameter $m$ and $\mu_r$ is the rth central moment, prove that*

$$m\,[{}^rC_1\,\mu_{r-1} + {}^rC_2\,\mu_{r-2} + \ldots + {}^rC_r\,\mu_0] = \mu_{r+1}.$$

**[Delhi Univ. B.Sc. (Stat. Hons.) 1990]**

**Solution** Since $X \sim P(m)$, its probability function is given by

$$p(x) = \frac{e^{-m} \cdot m^x}{x!}, \; x = 0, 1, 2, \ldots; m > 0$$

By definition,

$$\mu_{r+1} = E[X - E(X)]^{r+1} = E[X - m]^{r+1}$$

$$= \sum_{x=0}^{\infty} (x - m)^{r+1} p(x)$$

$$= \sum_{x=0}^{\infty} (x - m)^r (x - m) \frac{e^{-m} \cdot m^x}{x!}$$

$$= \sum_{x=0}^{\infty} \frac{x(x - m)^r e^{-m} m^x}{x!} - m \sum_{x=0}^{\infty} (x - m)^r \cdot \frac{e^{-m} m^x}{x!}$$

$$= \sum_{x=1}^{\infty} \frac{(x-m)^r e^{-m} m^x}{(x-1)!} - m\mu_r$$

$$= \sum_{y=0}^{\infty} \frac{(y-m+1)^r \cdot e^{-m} \cdot m^{y+1}}{y!} - m\mu_r, \qquad (x-1=y)$$

$$= m \cdot \sum_{y=0}^{\infty} (y-m+1)^r \cdot p(y) - m\mu_r$$

$$= m \sum_{y=0}^{\infty} \Big[ (y-m)^r + {}^rC_1 (y-m)^{r-1} + {}^rC_2 (y-m)^{r-2}$$

$$+ \dots + {}^rC_{r-1}(y-m) + 1 \Big] p(y) - m\mu_r$$

$$= m[\mu_r + {}^rC_1 \mu_{r-1} + {}^rC_2 \mu_{r-2} + \dots + {}^rC_r \mu_0] - m\mu_r$$

$$= m[{}^rC_1 \mu_{r-1} + {}^rC_2 \mu_{r-2} + \dots + {}^rC_r \mu_0].$$

**Example 7·42.** *If $X$ has a Poisson distribution with parameter $\lambda$, show that the distribution function of $X$ is given by*

$$F(x) = \frac{1}{\Gamma(x+1)} \int_{\lambda}^{\infty} e^{-t} t^x \, dt \, ; \; x = 0, 1, 2, \dots$$

**[Delhi Univ. M. Sc. (Stat) 1986]**

**Solution.** If $X$ is a Poisson variate, then

$$P(X=x) = \frac{e^{-\lambda} \lambda^x}{x!} \, ; \; x = 0, 1, 2, \dots \qquad (*)$$

Consider the incomplete gamma integral;

$$I_x = \frac{1}{x!} \int_{\lambda}^{\infty} e^{-t} t^x \, dt \, ; \quad (x \text{ is a positive integer})$$

$$= \left| -\frac{e^{-t} t^x}{x!} \right|_{\lambda}^{\infty} + \frac{1}{(x-1)!} \int_{\lambda}^{\infty} e^{-t} t^{x-1} \, dt$$

$$= \frac{e^{-\lambda} \lambda^x}{x!} + I_{x-1} \qquad (**)$$

which is a reduction formula for $I_x$.

Repeated applications of (**) gives

$$I_x = \frac{e^{-\lambda} \lambda^x}{x!} + \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} + \dots + \frac{e^{-\lambda} \lambda}{1!} + I_0$$

But $I_0 = \int_{\lambda}^{\infty} e^{-t} \, dt = \left| -e^{-t} \right|_{\lambda}^{\infty} = e^{-\lambda}$

$$\therefore \; I_x = e^{-\lambda} + \lambda e^{-\lambda} + \frac{\lambda^2 e^{-\lambda}}{2!} + \dots + \frac{\lambda^x}{x!} e^{-\lambda}$$

$$= P(X=0) + P(X=1) + \dots + P(X=x) \qquad [\text{From } (*)]$$

$$= P(X \leq x) = F(x)$$

where $F(\cdot)$ is the distribution function of the r.v. $X$.

$$\Rightarrow \qquad F(x) = \frac{1}{x!}\int_{\lambda}^{\infty} e^{-t}t^{x}\,dt = \frac{1}{\Gamma(x+1)}\int_{\lambda}^{\infty} e^{-t}t^{x}\,dt$$

($\because \Gamma(x+1) = x!$, since $x$ is a positive integer.)

**Remark.** This result is of great practical utility. It enables us to represent the cumulative Poisson probabilities (which are generally tedious to compute numerically) in terms of incomplete gamma integral, the values of which are tabulated for different values of $\lambda$ by Karl Pearson in his Tables of Incomplete $\Gamma$-Functions.

**7·3·10. Recurrence Formula for the Probabilities of Poisson Distribution.** *(Fitting of Poisson Distribution).* For a Poisson distribution with parameter $\lambda$, we have

$$p(x) = \frac{e^{-\lambda}\cdot\lambda^{x}}{x!}\,;\; x = 0, 1, 2, \ldots, \infty$$

and
$$P(x+1) = \frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!}\,;\; x = 0, 1, 2, \ldots, \infty$$

$$\therefore \qquad \frac{p(x+1)}{p(x)} = \frac{\lambda}{(x+1)} \Rightarrow p(x+1) = \frac{\lambda}{x+1}p(x) \qquad \ldots(17\cdot20)$$

which is the required recurrence formula.

This formula provides us a very convenient method of graduating the given data by a Poisson distribution. The only probability we need to calculate is $p(0)$ which is given by $p(0) = e^{-\lambda}$, where $\lambda$ is estimated from the given data. The other probabilities, *viz.*, $p(1), p(2)\ldots$. can now be easily obtained as explained below:

$$p(1) = [p(x+1)]_{x=0} = \left[\frac{\lambda}{x+1}\right]_{x=0} p(0),$$

$$p(2) = [p(x+1)]_{x=1} = \left[\frac{\lambda}{x+1}\right]_{x=1} p(1),$$

$$p(3) = [p(x+1)]_{x=2} = \left[\frac{\lambda}{x+1}\right]_{x=2} p(2),$$

and so on.

**Example 7·43.** *After correcting 50 pages of the proof of a book, the proof reader finds that there are, on the average, 2 errors per 5 pages. How many pages would one expect to find with 0, 1, 2, 3 and 4 errors, in 1000 pages of the first print of the book ? (Given that $e^{-0\cdot4} = 0\cdot6703$)*

**[Nagpur Univ. M.A. (Eco.), 1989]**

**Solution.** Let the random variable $X$ denote the number of errors per page. Then the mean number of errors per page is given by :

$$\lambda = 2/5 = 0{\cdot}4$$

Using Poisson probability law, probability of $x$ errors per page is given by:

$$P(X = x) = p(x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-0\cdot4}(0{\cdot}4)^x}{x!}; x = 0, 1, 2, \ldots$$

Expected number of pages with $x$ errors per page in a book of 1000 pages are :

$$1000 \times P(X = x) = 1000 \times \frac{e^{-0\cdot4}(0{\cdot}4)^x}{x!}; x = 0, 1, 2, \ldots$$

Using the recurrence formula ( 17·20 ), various probabilities can be easily calculated as shown in the following table.

| No. of errors per page ($X$) | Probability $p(x)$ | Expected number of pages $1000\,p(x)$ |
|---|---|---|
| 0 | $p(0) = e^{-0\cdot4} = 0{\cdot}6703$ | $670{\cdot}3 \simeq 670$ |
| 1 | $p(1) = \frac{0\cdot4}{0+1}\,p(0) = 0{\cdot}26812$ | $268{\cdot}12 \simeq 268$ |
| 2 | $p(2) = \frac{0\cdot4}{1+1}\,p(1) = 0{\cdot}053624$ | $53{\cdot}624 \simeq 54$ |
| 3 | $p(3) = \frac{0\cdot4}{2+1}\,p(2) = 0{\cdot}0071298$ | $7{\cdot}1298 \simeq 7$ |
| 4 | $p(4) = \frac{0\cdot4}{3+1}\,p(3) = 0{\cdot}00071298$ | $0{\cdot}71298 \simeq 1$ |

**Example 7·44.** *Fit a Poisson distribution to the following data which gives the number of doddens in a sample of clover seeds.*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *No. of doddens:* ($x$) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| *Observed frequency:* ($f$) | 56 | 156 | 132 | 92 | 37 | 22 | 4 | 0 | 1 |

**Solution.**

$$\text{Mean} = \frac{1}{N}\sum fx = \frac{986}{500} = 1{\cdot}972$$

Taking the mean of the given distribution as the mean of the Poisson distribution we want to fit, we get $\lambda = 1{\cdot}972$,

and $$p(x) = \frac{e^{-\lambda}\cdot\lambda^x}{x!}; \; x = 0, 1, 2, \ldots, \infty$$

$$p(0) = e^{-\lambda} = e^{-1\cdot972}$$

$\therefore \quad \log_{10} p(0) = -1\cdot972 \log_{10} e = -1\cdot972 \times 0\cdot43429$

$= -0\cdot856419 = \bar{1}\cdot143581$

$\therefore \quad p(0) = 0\cdot1392$

Using the recurrence formula (17·20) the various probabilities, *viz.*, $p(1), p(2), \ldots$, can be easily calculated as shown in the following table :

| $x$ | $\dfrac{\lambda}{x+1}$ | $p(x)$ | Expected frequency $N.p(x)$ |
|---|---|---|---|
| 0 | 1·972 | 0·13920 | 69·6000 |
| 1 | 0·986 | 0·27455 | 137·2512 |
| 2 | 0·657 | 0·27006 | 135·3296 |
| 3 | 0·493 | 0·17793 | 88·9566 |
| 4 | 0·394 | 0·10964 | 43·8556 |
| 5 | 0·328 | 0·03459 | 17·2966 |
| 6 | 0·281 | 0·01137 | 5·6846 |
| 7 | 0·247 | 0·00320 | 1·6013 |
| 8 | 0·219 | 0·00078 | 0·3942 |

Since frequencies are always integers, therefore by converting them to nearest integers, we get

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Observed frequency :* | 56 | 156 | 132 | 92 | 37 | 22 | 4 | 0 | 1 |
| *Expected frequency :* | 70 | 137 | 135 | 89 | 44 | 17 | 6 | 2 | 0 |

**Remark.** In rounding the figures to the nearest integer it has to be kept in mind that the total of the observed and the expected frequencies should be same.

## EXERCISE 7 (b)

1. (*a*) Derrive Poisson distribution as a limiting form of a binomial distribution. **[Madras Univ. B. E., Dec. 1991]**

Hence find $\beta_1$ and $\beta_2$ of the distribution.

Give some examples of the occurrence of Poisson distribution in different fields.

(*b*) State and prove the reproductive property of the Poisson distribution. Show that the mean and variance of the Poisson distribution are equal.

Find the mode of the Poisson distribution with mean value 5.

(*c*) Prove that under certain conditions to be stated by you, the number of telephone calls on a trunkline in a given interval of time has a Poisson distribution.

**[Calcutta Univ. B.Sc. (Maths Hons.), 1989]**

(*d*) Show that for a Poisson distribution, the coefficient of variation is the reciprocal of the standard deviation.

2. (*a*) If two independent variables $X_1$ and $X_2$ have Poisson distribution with means $\lambda_1$ and $\lambda_2$ respectively, then show that their sum $X_1 + X_2$ is a Poisson variate with mean $\lambda_1 + \lambda_2$.

Does the difference of two independent Poisson variates follow a Poisson distribution ? Give reasons. [Sri Venketeswara Univ. B.Sc., 1991]

(*b*) Prove that the sum of two independent Poisson variates is a Poisson variate. Is the result true for the difference also ? Give reasons.

[Delhi Univ. B.Sc. (Stat. Hons.) 1989]

(*c*) If $X_1, X_2, ..., X_k$ are independent random variables following the Poisson law with parameter $m_1, m_2, ..., m_k$ respectively, show that $\sum_{i=1}^{k} X_i$ follows the Poisson law with parameter $\sum_{i=1}^{k} m_i$ [Madras Univ. B. E., 1993]

3. (*a*) Prove the recurrence relation between the moments of Poisson distribution

$$\mu_{r+1} = \lambda \left( r\mu_{r-1} + \frac{d\mu_r}{d\lambda} \right), \text{ where } \mu_r = \sum_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^j}{j!}(j-\lambda)^r$$

where $\mu_r$ is the $r$th moment about the mean $\lambda$. Hence obtain the skewness and kurtosis of Poisson distribution.

[Delhi Univ. B. Sc. (Stat. Hons.) 1989,' 86; Utkal Univ. B. Sc. 1993]

(*b*) Let $X$ have a Poisson distribution with parameter $\lambda > 0$. If $r$ is a non-negative integer and if $\mu_r' = E(X^r)$, prove that

$$\mu'_{r+1} = \lambda \left( \mu'_r + \frac{d\mu'_r}{d\lambda} \right)$$

[Madras Univ. B. Sc. Nov. 1988]

4. What do you understand by (*i*) cumulants, (*ii*) cumulative function. Obtain the cumulative function of a Poisson distribution with parameter $\lambda$. Hence or otherwise show that for a Poisson distribution with parameter $\lambda$, all the cumulants are $\lambda$.

5. For the Poisson distribution with parameter $\lambda$, show that the $r$th factorial moment $\mu'_{(r)}$ is given by $\mu'_{(r)} = \lambda^r$

Show further that $\mu_{(2)} = \lambda$, $\mu_{(3)} = -2\lambda$ and $\mu_{(4)} = 3\lambda(\lambda + 2)$

6. (*a*) If $X$ and $Y$ are independent *r.v. s.*' so that $X \sim P(\lambda)$ and $X + Y \sim P(\lambda + \mu)$, find the distribution of $Y$. [Ans. $Y \sim P(\mu)$]

(*b*) If $X \sim P(\lambda)$, find

(*i*) Karl Pearson's coefficient of skewness

(*ii*) Moment measure of skewness.

Is Poisson distribution positively skewed or negatively skewed ?

**7.** (*a*) It is known that the probability that an item produced by a certain machine will be defective is 0·01. By applying Poisson's approximation, show that the probability that random sample of 100 items selected at random from the total output will contain no more than one defective item is $2/e$.

(*b*) The probability of success in a trail is known to be $10^{-4}$. It is possible to repeat the trial independently any desired number of times. Do you think that the number of successes in a series of trials, if the number of trials in the series increases indefinitely, will tend to follow a Poisson distribution ? Give your reasons.

(*c*) The probability of getting no misprint in a page of a book is $e^{-4}$. What is the probability that a page contains more than 2 misprints ? [State the assumptions you make in solving this problem.] **[Bombay Univ. B.Sc., 1989]**

**8.** In a certain factory turning out optical lenses, there is a small chance 1/500 for any lens to be defective. The lenses are supplied in a packet of 10. Use Poisson distribution to calculate the approximate number of packets containing no defective, one defective, two defective and three defective lenses in a consignment of 20,000 packets.

**Ans.** 19604, 392, 4 and 0 packets.

**9.** Red blood cell deficiency may be determined by examining a specimen of the blood under a microscope. Suppose a certain small fixed volume contains on the average 20 red cells for normal persons. Using Poisson distribution, obtain the probability that a specimen from a normal person will contain less than 15 red cells.

**Ans.** $$\sum_{x=0}^{14} \{e^{-20} (20)^x/x\,!\}$$

**10.** Assuming that the chance of a traffic accident in a day in a street of Delhi is 0·001, on how many days out of a trial of 1,000 days can we expect :

(*i*) no accident

(*ii*) more than three accidents, if there are 1,000 such streets in the whole city ?

**11.** Patients arrive randomly and independently at a doctor's surgery from 8·0 A.M. at an average rate of one in five minutes.The waiting room holds 2 persons. What is the probability that the room will be full when the doctor arrives at 9·0 A.M. (Estimate the probability to an accuracy of 5 per cent.)

**Ans.** 53·84 %

**12.** An office switchboard receives telephone calls at the rate of 3 calls per minute on an average. What is the probability of receiving (*i*) no calls in a one-minute interval, (*ii*) at the most 3 calls in a 5-minute interval ?

**Ans.** (*i*) 0·0323, (*ii*) 0

**13.** A hospital switchboard receives an average of 4 emergency calls in a 10-minute interval. What is the probability that (*i*) there are at the most 2 emergency calls in a 10-minute interval, (*ii*) there are exactly 3 emergency calls in a 10-minute interval ?

**Ans.** (*i*) $13^{-4}$, (*ii*) $(32/3)\, e^{-4}$

**14.** (*a*) A distributor of bean seeds determines from extensive tests that 5% of large batch of seeds will not germinate. He sells the seeds in packets of 200 and guarantees 90% germination. Determine the probability that a particular packet will violate the guarantee.

**Ans.** $1 - \sum_{r=0}^{10} (e^{-10}\, 10^r/r!)$

(*b*) In an automatic telephone exchange the probability that any one call is wrongly connected is 0·001. What is the minimum number of independent calls required to ensure a probability of 0·90, that at least one call is wrongly connected?

**15.** (*a*) Fit a Poisson distribution to the following data with respect to the number of red blood corpuscles ($x$) per cell :

| $x$ : | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Number of cells $f$ : | 142 | 156 | 69 | 27 | 5 | 1 |

(*b*) Data was collected over a period of 10 years, showing number of deaths from horse kicks in each of the 20 army corps. From the 200 corps-years, the distribution of deaths was as follows :

| No. of deaths : | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Frequency : | 122 | 60 | 15 | 2 | 1 |

Graduate the data by Poisson distribution and calculate the theoretical frequencies.

| Given | $e^{-m}$ : | 0·6703 | 0·6065 | 0·5488 | 0·4966 |
|---|---|---|---|---|---|
| | $m$ : | 0·4 | 0·5 | 0·6 | 0·7 |

(c) Fit a Poisson distribution to the following data and calculate the expected frequencies :-

| $x$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| $f$ : | 71 | 112 | 117 | 57 | 27 | 11 | 3 | 1 | 1 |

**16.** (*a*) If $X$ is the number of occurrences of the Poisson variate with mean $\lambda$ ; show that : $P(X \geq n) - P(X \geq n+1) = P(X = n)$

(*b*) Suppose that $X$ has a Poisson distribution. If

$$P(X = 2) = \tfrac{2}{3}\, P(X = 1).$$

Evaluate (*i*) $P(X = 0)$ and (*ii*) $P(X = 3)$ [Ans. (*i*) 0·264.]

(c) If $X$ has a Poisson distribution such that

$P(X = 1) = P(X = 2)$, find $P(X = 4)$. [Ans 0·09]

(c) If a Poisson variate $X$ is such that

$$P(X = 1) = 2\,P(X = 2),$$

find $P(X = 0)$, mean and the variance.

(*d*) If for a Poisson variate $X$, $E(X^2) = 6$, what is $E(X)$ ?

(*e*) If $X$ and $Y$ are independent Poisson variates having means 1 and 3 respectively, find the variance of $3X + Y$.

**17.** Show that for a Poisson distribution

(*i*) $M \sigma \gamma_1 \gamma_2 = 1$, (*ii*) $\beta_1^{1/2} (\beta_2 - 3) \mu_1' \sigma = 1$

**18.** Show that the function which generates the central moments of the Poisson distribution with parameter $\lambda$ is

$$M(t) = \exp\{\lambda(e^t - 1 - t)\}$$

Show that it satisfies the equation

$$\frac{dM(t)}{dt} = \lambda t\, M(t) + \lambda \frac{dM(t)}{d\lambda}$$

**19.** (*a*) The random variable $X$ has p.d.f.

$$f(x) = e^{-\theta} \frac{\theta^x}{x!};\ x = 0, 1, 2, \ldots$$

$$= 0, \text{ elsewhere}$$

Find the m.g.f. of $Y = 2X - 1$ and Var $(Y)$.

(*b*) Identify the distribution with the following mgf's :

$$M_X(t) = (0{\cdot}3 + 0{\cdot}7\, e^t)^{10}$$

$$M_Y(t) = \exp[3(e^t - 1)]$$

**Ans.** $X \sim B(10, 0{\cdot}7)$, $Y \sim P(3)$.

**20.** If $X$ has Poisson distribution with parameter $\lambda$, then

$$P[X \text{ is even}] = \frac{1}{2}\left[1 + e^{-2\lambda}\right]$$

**[Delhi Univ. B. Sc. (Stat. Hons.) 1991]**

**21.** (*a*) The m.g.f. of a r.v. is $X$ is $\exp[4(e^t - 1)]$. Show that

$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0{\cdot}931$

**Hint.** $X \sim P(\lambda = 4)$;

Required Probability. $= P(0 < X < 8) = P(1 \le X \le 7) = 0{\cdot}931$

(*b*) If $X \sim P(\lambda = 100)$, use Chebychev's inequality to determine a lower bound for $P(75 < X < 125)$ [Ans. 0·84]

**22.** If $X \sim P(m)$, show that $E|X - 1| = m - 1 + 2e^{-m}$

**[Delhi Univ. B. Sc. (Maths. Hons.), 1983]**

**Hint.** $E|X - 1| = \sum_{x=0}^{\infty} |x - 1|\, e^{-m} m^x / x! = e^{-m} + \sum_{x=2}^{\infty} \frac{(x-1)}{x!} \cdot e^{-m} m^x$

$$= e^{-m} + e^{-m} \cdot \sum_{x=2}^{\infty} m^x \left[\frac{1}{(x-1)!} - \frac{1}{x!}\right]$$

**23.** If $X \sim P(\lambda)$ and $Y|X = x \sim (B(x, p)$, then prove that $Y \sim P(\lambda p)$.

**24.** If the chances of 0, 1, 2, 3... events from one source are given by a Poisson distribution of mean $m_1$ and the chances of 0, 1, 2, 3,... events from another source by a Poisson distribution of mean $m_2$, show that the chances of 0, 1, 2, 3,... events from either source are given by

$$e^{-(m_1+m_2)}\left\{1,\ (m_1 + m_2),\ \frac{(m_1 + m_2)^2}{2!}, \dots\right\}.$$

Show that the sum of any finite number of Poisson variates is itself a Poisson variate with mean equal to the sum of separate means.

**25.** $X$ is a Poisson variate with mean $\lambda$.

Show that $E(X^2) = \lambda E(X + 1)$.

If $\lambda = 1$, show that $E|X - 1| = \dfrac{2}{e}$

**26.** Show that the mean deviation about mean for Poisson distribution

$$p(x) = \frac{e^{-m} m^x}{x!};\ x = 0, 1, 2, \dots$$

is $(2\mu) \cdot \dfrac{e^{-m} \cdot m^{\mu}}{\mu!}$

where $\mu$ is the greatest integer contained in $(m + 1)$.

[Delhi Univ. B. Sc. (Stat. Hons.), 1988,' 84]

**27.** Let $X$, $Y$ be independent Poisson variates. The variance of $X + Y$ is 9 and

$$P(X = 3 | X + Y = 6) = 5/54$$

Find the mean of $X$. [Ans. $\frac{1}{2}(9 \pm 3\sqrt{3})$ *i.e.* 1·902 *or* 7·098]

**28.** If $X$ is a Poisson variate with paramter $m$, show that

$$P(X < r) < \frac{m^r}{r!};\ r = 0, 1, 2, \dots$$

Deduce that $E(X) < e^m$. [Delhi Univ. B.Sc. (Maths. Hons.), 1989]

**29.** (*a*) The characteristic function of a variate $X$ is

$$\varphi_X(t) = \left(\frac{1}{3} + \frac{2}{3}e^{it}\right)^6 \cdot \left[\exp\left\{-3(1 - e^{it})\right\}\right].$$

Recognise the variate.

[Burdwan Univ. B. Sc. (Maths. Hons.) 1989)

**Hint.** $X = U + V$, where $U \sim B\left(6, \frac{2}{3}\right)$ and $V \sim P(3)$ are independent *r.v.'s*

(*b*) Identify the variates $X$ and $Y$ where :

$M_X(t) = (1/27)(1 + 2e^t)^3 \cdot \exp\left[3(e^t - 1)\right]$

$$M_Y(t) = (1/32)(1+e^t)^5 \cdot \exp\left[-2(1-e^t)\right]$$

**[Delhi Univ. B. Sc. (Stat. Hons.), 1987, 84]**

**Ans.** $X = U + V$ ; $U \sim B(n = 3, p = 2/3)$ and $V \sim P(\lambda = 3)$ are independent.

$Y = U_1 + V_1$ ; $U_1 \sim B(n = 5, p = 1/2)$ and $V_1 \sim P(\lambda = 2)$ are independent.

**30.** If $X$ and $Y$ are correlated variates each having Poisson distribution, show that $X + Y$ cannot be a Poisson variate

**[Delhi Univ. B. Sc. (Maths Hons.), 1988; Poona Univ. B.Sc., 1989]**

**Hint.** Note that for Poisson variate mean and variance are equal. Let $X \sim P(\lambda)$, $Y \sim P(\mu)$ ; $(X, Y)$ correlated.

$$\therefore E(X+Y) = E(X) + E(Y) = \lambda + \mu$$

$$\begin{aligned}\text{Var}(X+Y) &= \text{Var}\,X + \text{Var}\,Y + 2\,\text{Cov}(X,Y)\\ &= \lambda + \mu + 2\rho\sqrt{\lambda\mu},\ (\rho \neq 0)\end{aligned}$$

Since $E(X+Y) \neq \text{Var}(X+Y)$ ; $X + Y$ cannot be a Poisson variate.

**31.** Let $X, Y, Z$ be independent Poisson variates with parameters $a, b$ and $c$ respectively. Obtain :

(*i*) m.g.f. of $X + 2Y + 3Z$,

(*ii*) Conditional expectation of $X$ given $X + Y + Z = n$

**(Indian Civil Services, 1985)**

**Hint.** $M_{X+2Y+3Z}(t) = M_X(t) \cdot M_Y(2t) \cdot M_Z(3t)$

$$= \exp\left[a\left(e^t - 1\right) + b\left(e^{2t} - 1\right) + c\left(e^{3t} - 1\right)\right]$$

$$P(X = x \mid X + Y + Z = n) = \frac{P(X = x \cap X + Y + Z = n)}{P(X + Y + Z = n)}$$

$$= \frac{P(X = x)P(Y + Z = n - x)}{P(X + Y + Z = n)} \quad (\because X, Y, Z \text{ are indep.})$$

$$= \frac{e^{-a}\cdot a^x}{x!} \times \frac{e^{-(b+c)}\cdot(b+c)^{n-x}}{(n-x)!} \times \left[\frac{n!}{e^{-(a+b+c)}\cdot(a+b+c)^n}\right]$$

$$= \frac{n!}{x!(n-x)!}\left(\frac{a}{a+b+c}\right)^x \cdot \left(\frac{b+c}{a+b+c}\right)^{n}$$

$$\Rightarrow \quad X \mid (X + Y + Z = n) \sim B\{n,\ p = a/(a+b+c)\}$$

$$\Rightarrow \quad E[X \mid X + Y + Z = n] = np = \frac{na}{a+b+c}$$

**32.** The joint density of r.v.'s $X$ and $Y$ is :

$$f(x,y) = e^{-2} / [x!\,(y-x)!] ;\ y = 0, 1, 2, \ldots ;\ x = 0, 1, 2, \ldots, y.$$

Find the m.g.f. $M(t_1, t_2)$ of $(X, Y)$ and correlation coefficient between $X$ and $Y$. Show that the marginal distributions of $X$ and $Y$ are Poisson.

**Hint.** $M(t_1, t_2) = \sum\limits_{y=0}^{\infty} \sum\limits_{x=0}^{y} e^{t_1 x + t_2 y} \times \left[ \dfrac{e^{-2}}{x!(y-x)!} \right]$

$$= e^{-2} \sum_{y=0}^{\infty} \left[ \frac{e^{t_2 y}}{y!} \left\{ \sum_{x=0}^{y} {}^yC_x \cdot (e^{t_1})^x \right\} \right]$$

$$= e^{-2} \sum_{y=0}^{\infty} \left\{ \left[ e^{t_2} \left( 1 + e^{t_1} \right) \right]^y / y! \right\}$$

$$= e^{-2} \cdot \exp \left\{ e^{t_2} ( 1 + e^{t_1} ) \right\}$$

$$M(t_1, 0) = \exp\left[ 2(e^{t_1} - 1) \right] \Rightarrow X \sim P(\lambda = 1)$$

$$M(0, t_2) = \exp\left[ 2(e^{t_2} - 1) \right] \Rightarrow Y \sim P(\mu = 2)$$

Observe $M(t_1, t_2) \neq M(t_1, 0) \times M(0, t_2) \Rightarrow X$ and $Y$ are not independent.

$E(X) = 1, \quad \text{Var}(X) = 1; \quad E(Y) = 2 = \text{Var}\, Y.$

$$E(XY) = \left| \frac{\partial^2 M(t_1, t_2)}{\partial t_1 \partial t_2} \right|_{t_1 = t_2 = 0} = 3$$

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{3 - 1 \times 2}{1 \times \sqrt{2}} = 1/\sqrt{2}.$$

33. The joint p.g.f. of the r.v.'s $X$ and $Y$ is given by :

$$P(s_1, s_2) = \exp\left[ a(s_1 - 1) + b(s_2 - 1) + c(s_1 - 1)(s_2 - 1) \right],$$

a,b,c, are all positive. Find $\rho(X, Y)$

**Hint.** $P_X(s_1) = P(s_1, 1) = \exp\left[ a(s_1 - 1) \right] \Rightarrow X \sim P(a)$

$$P_Y(s_2) = P(1, s_2) = \exp\left[ b(s_2 - 1) \right] \Rightarrow Y \sim P(b)$$

$$E(XY) = \left( \frac{\partial^2 P(s_1, s_2)}{\partial s_1 \partial s_2} \right)_{s_1 = s_2 = 1} = c + ab.$$

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{(c + ab) - ab}{\sqrt{a}\sqrt{b}} = \frac{c}{\sqrt{ab}}$$

34. An insurance company issues only two types of policy, household and motor. It has carried out an investigation into the experience of a group of policyholders who held one of each type of policy over a particular period and it has discovered that within that group and over that period the mean number of claims per household policy was 0·3 and the mean number of claims per motor policy was 0·8. Assume that the number of claims under each type of policy is independent of the number of claims under the other type of policy and that each can be represented by a Poisson distribution.

(*a*) If the number of claims per policyholder is the sum of the number of claims under each of his two policies, state with reasons how the number of claims per policyholder, within that group and over that period is distributed, and (*b*) Calculate to the nearest whole number, the percentage of policyholders within that group and over that period who made more household claims than motor claims.

**Hint.** Household claim, $X \sim P(\cdot 3)$ and Motor claim, $Y \sim P(\cdot 8)$

$$\text{Required Probability} = P(X > Y) = \sum_{r=0}^{\infty}\left[\sum_{s=0}^{\infty} P(Y = r \cap X = r + s)\right]$$

$$= \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\left[P(Y=r)P(X=r+s)\right] = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{e^{-\cdot 8}(\cdot 8)^r}{r!} \times \frac{e^{-\cdot 3}(\cdot 3)^{r+s}}{(r+s)!}$$

$$= e^{-\cdot 8}\, e^{-\cdot 3}\sum_{r=0}^{\infty}\left[\frac{(\cdot 8)^r}{r!}\sum_{s=0}^{\infty}\left\{\frac{(\cdot 3)^{r+s}}{(r+s)!}\right\}\right]$$

$$= \sum_{r=0}^{\infty}\left[\frac{e^{-\cdot 8}(\cdot 8)^r}{r!}\, e^{-\cdot 3}\left\{e^{\cdot 3} - \left(1 + \cdot 3 + \frac{(\cdot 3)^2}{2!} + \dots + \frac{(\cdot 3)^{r-1}}{(r-1)!}\right)\right\}\right]$$

$$= 1 - e^{-\cdot 8} e^{-\cdot 3}\left[\left\{\frac{\cdot 8}{1} + \frac{(\cdot 8)^2}{2!}(1 + \cdot 3) + \frac{(\cdot 8)^3}{3!}\left(1 + \cdot 3 + \frac{\cdot 09}{2}\right)\right.\right.$$

$$\left. + \frac{(\cdot 8)^4}{4!}\left(1 + \cdot 3 + \frac{\cdot 09}{2} + \frac{\cdot 027}{3!}\right) + \dots\right]$$

35. (*i*) An event occurs instantaneously and is equally likely to occur at any instant. There is no limit on the number of occurrneces that may happen in any interval of time, but the expected number in a given time interval is $T$. Prove that the probability of the event occurring exactly $r$ times in an interval of the same duration is $(T^r\, e^{-T})/r\,!$.

(*ii*) An insurance company which writes only fire and accident business defines a major claim as one which costs at least Rs. 50,000 for an accident claim or Rs. 100,000 for a fire claim. Any excess over these amounts is paid by reinsurers and hence every major claim is recorded at a cost of Rs. 50,000 or Rs. 100,000 respectively. The company divides the year into equal monthly accounting periods and a report is produced of the recorded cost of major claims. The expected number of major accident claims is 0·2 per month and of major fire claims 0·5 per month. Calculate the probability that in a particular month the recorded cost of major claims is Rs. 2,00,000 or more.

36. (*a*) The number of aeroplanes arriving at an airport in a 30 minute interval obeys the Poisson law with mean 25. Use Chebychev's inequality to find the least chance, that the number of planes to arrive within a given 30 minutes interval will be between 15 and 35. **[Sri Venketeswara U. B.Sc. 1992]**

(*b*) Suppose that the number of motor cars arriving in a certain parking lot in any 15 minutes period obeys a Poisson probability law with mean 80. Use Chebychev's inequality to determine a lower bound for the probability that the number of motor cars arriving in a given 15 minute period will be between 60 and 100. [Madras U. B.Sc. Nov. 1991]

**7·4. Negative Binomial Distribution.** The equality of the mean and varaince is an important characteristic of the Poisson distribution, whereas for the binomial distribution the mean is always greater than the variance. Occasionally, however, observable phenomena give rise to empirical discrete distributions which show a variance larger than the mean. Some of the commonest examples of such behaviour are the frequency distributions of plant density obtained by quadrant sampling when the clustering of plants makes the simple Poisson model inapplicable. It has been shown by different investigators that in such cases the negative binomial distribution provides an excellent model because this distribution has a variance larger than the mean. Bacterial clustering (or contagion), *e.g.*, deaths of insects, number of insect bites leads to the negative binomial distribution and the distribution also arises in inverse sampling from a binomial population or as a weighted average of Poisson distribution. This important probability distribution is sometimes also referred to as the Pascal distribution after the French mathematician Blaise Pascal (1623-1662), but there seems to be no historical justification. The negative binomial distribution can be derived from empirical considerations in many ways. Here we consider the Binomial probability situation with some modifications.

Suppose we have a succession of $n$ Bernoulli trails. We assume that (*i*) the trials are independent, (*ii*) the probability of success '$p$' in a trial remains constant from trial to trial.

Let $f(x\,;r,p)$ denote the probability that there are $x$ failures preceding the $r$th success in $x+r$ trials.

Now, the last trial must be a success, whose probability is $p$. In the remaining $(x+r-1)$ trials we must have $(r-1)$ successes whose probability is given by

$$\binom{x+r-1}{r-1}p^{r-1}q^x$$

Therefore by compound probability theorem, $f(x\,;r,p)$ is given by the product of these two probabilities, *i.e.*,

$$\binom{x+r-1}{r-1}p^{r-1}q^x\cdot p=\binom{x+r-1}{r-1}p^r q^x$$

**Definition.** *A random variable X is said to follow a negative binomial distribution if its probability mass function is given by*

$$p(x)=P(X=x)=\binom{x+r-1}{r-1}p^r q^x\,;x=0,1,2,\ldots$$
$$=0,\ \textit{otherwise} \qquad \ldots(7\cdot21)$$

Also

$$\binom{x+r-1}{r-1}=\binom{x+r-1}{x} \qquad \left[\because\binom{n}{r}=\binom{n}{n-r}\right]$$

$$= \frac{(x+r-1)(x+r-2)\ldots(r+1)r}{x!}$$

$$= \frac{(-1)^x(-r)(-r-1)\ldots(-r-x+2)(-r-x+1)}{x!}$$

$$= (-1)^x \binom{-r}{x}$$

$$\therefore \qquad p(x) = \begin{cases} \binom{-r}{x} p^r (-q)^x ; x = 0, 1, 2, \ldots \\ 0, \text{ otherwise} \end{cases} \qquad \ldots(7\cdot21\ a)$$

which is the $(x+1)^{th}$ term in the expansion of $p^r(1-q)^{-r}$, a binomial expansion with a negative index. Hence the distribution is known as negative binomial distribution. Also

$$\sum_{x=0}^{\infty} p(x) = p^r \sum_{x=0}^{\infty} \binom{-r}{x} (-q)^x = p^r \times (1-q)^{-r} = 1$$

Therefore $p(x)$ represents the probability function and the discrete variable which follows this probability function is called the negative binomial variate.

If $\quad p = \frac{1}{Q}$ and $q = \frac{P}{Q}$ so that $Q - P = 1$, $\quad (\because p + q = 1)$

$$\text{then} \quad p(x) = \begin{cases} \binom{-r}{x} Q^{-r} \left(-\frac{P}{Q}\right)^x ; x = 0, 1, 2, \ldots \\ 0, \text{ otherwise} \end{cases} \qquad \ldots(7\cdot21\ b)$$

This is the general term in the negative binomial expansion $(Q - P)^{-r}$.

**Remarks. 1.** $p(x)$ in (7·21) or (7·21*a*) is also sometimes written as $f(x; r, p)$.

**2. Some Important Deductions.**

(*a*) **Geometric Distribution.** If we take $r = 1$ in (7·21), we have

$$p(x) = q^x p ; \ x = 0, 1, 2, \ldots$$

which is the probability function of geometric distribution (*c.f.* § 7·5 page 7·83).

Hence negative binomial distribution may be regarded as the generalisation of geometric distribution.

(*b*) **Pascal's Distribution.** The negative binomial distribution (7·21 *a*) when regarded as one having two parameters $p$ and $r$ is known as *Pascal's distribution.*

(*c*) **Polya's Distribution.** If we take

$r = \frac{1}{\beta}, \ p = \frac{1}{1 + \beta\mu}, \ q = 1 - p = \frac{\beta\mu}{1 + \beta\mu}$ in (7·21 *a*), we get

$$p(x) = \frac{r(r+1)(r+2)\ldots(r+x-1)}{x!} \cdot p^r \cdot q^x$$

$$= \frac{(1+\beta)(1+2\beta)\dots[1+\beta(x-1)]}{x!}\left(\frac{1}{1+\beta\mu}\right)^{1/\beta}\left(\frac{\mu}{1+\beta\mu}\right)^{x}$$

$$(x = 0, 1, 2, \dots) \qquad \dots(7{\cdot}21\ c)$$

which is known as *Polya's distribution* with two parameters, $\beta$ and $\mu$.

**(d) Second Form of Geometric Distribution.** Taking $\beta = 1$ in Polya's distribution (7·21 *c*), we get

$$p(x) = \left(\frac{1}{1+\mu}\right)\left(\frac{\mu}{1+\mu}\right)^{x}; \ x = 0, 1, 2, \dots \qquad \dots(7{\cdot}21\ d)$$

which is geometric distribution (c.f. § 7·5) with

$$p = \frac{1}{1+\mu}, \ q = 1 - p = \frac{\mu}{1+\mu}$$

**7·4·1. Moment Generating Function of Negative Binomial Distribution.**

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} p(x)$$

$$= \sum_{x=0}^{\infty} \binom{-r}{x} Q^{-r}\left(-\frac{Pe^t}{Q}\right)^{x}$$

$$= (Q - Pe^t)^{-r} \qquad \dots(7{\cdot}22)$$

$$\mu_1' = \left(\frac{d}{dt} M(t)\right)_{t=0} = \left[-r(-Pe^t)(Q-Pe^t)^{-r-1}\right]_{t=0}$$

$$= rP$$

$\therefore$ Mean of the negative binomial distribution is $rP$.

$$\mu_2' = \left(\frac{d^2}{dt^2} M(t)\right)_{t=0} \qquad \dots(7{\cdot}22\ a)$$

$$= \left(rPe^t(Q-Pe^t)^{-r-1} + (-r-1)rPe^t(Q-Pe^t)^{-r-2}(-Pe^t)\right)_{t=0}$$

$$= rP + r(r+1)P^2$$

$$\therefore \quad \mu_2 = \mu_2' - \mu_1'^2 = r(r+1)P^2 + rP - r^2p^2 = rPQ \qquad \dots(7{\cdot}22\ b)$$

As $Q > 1$, $rP < rPQ$, *i.e., Mean < Variance, which is a distinguishing feature of this distribution.*

**7·4·2. Cumulants of Negative Binomial Distribution.**

$$K_X(t) = \log M_X(t) = -r\log(Q - Pe^t)$$

$$= -r\log\left[Q - P\left(1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \dots\right)\right]$$

$$= -r\log\left[1 - P\left(t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \dots\right)\right]$$

$$(\because Q - P = 1)$$

Proceeding as in § 7·2·8, we will get (on replacing $n$ with $-r$ and $p$ with $-P$).

$$\left.\begin{aligned} \text{Mean} &= \kappa_1 = rP \\ \mu_2 &= \kappa_2 = r\,P(1+P) = rPQ \end{aligned}\right\} \qquad \ldots(7\cdot23)$$

$$\mu_3 = \kappa_3 = rP(1+3P+2P^2) = rP(1+P)(1+2P) = rPQ(Q+P).$$

$$\kappa_4 = rP(1+P)(1+6P+6P^2) = rPQ(1+6PQ)$$

$$\therefore \quad \mu_4 = \kappa_4 + 3\kappa_2^2 = rPQ[1+3PQ(r+2)]$$

Since $Q = 1/p$, $P = qQ = q/p$, we have in terms of $p$ and $q$,

Mean $= rq/p$, Variance $= rq/p^2$, $\mu_3 = rq(1+q)/p^3$

$$\mu_4 = rq\,[p^2 + 3q(r+2)]/p^4$$

$$\therefore \quad \beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(1+q)^2}{rq}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{p^2 + 3q(r+2)}{rq} \qquad \ldots(7\cdot23\,a)$$

$$\therefore \quad \gamma_1 = \sqrt{\beta_1} = (1+q)/\sqrt{rq}$$

$$\gamma_2 = \beta_2 - 3 = (p^2 + 6q)/rq$$

**7·4·3. Poisson Distribution as a Limiting Case of the Negative Binomial Distribution.** Negative binomial distribution tends to Poisson distribution as $P \to 0$, $r \to \infty$ such that $rP = \lambda$ (finite). Proceeding to the limits, we get

$$\lim p(x) = \lim \binom{x+r-1}{r-1} p^r q^x$$

$$= \lim \binom{x+r-1}{x} Q^{-r} \left(\frac{P}{Q}\right)^x$$

$$= \lim_{r\to\infty} \frac{(x+r-1)(x+r-2)\ldots(r+1)r}{x!} (1+P)^{-r} \left(\frac{P}{1+P}\right)^x$$

$$= \lim_{r\to\infty} \left[\frac{1}{x!}\left(1+\frac{x-1}{r}\right)\left(1+\frac{x-2}{r}\right) \ldots \left(1+\frac{1}{r}\right).1.r^x(1+P)^{-r}\left(\frac{P}{1+P}\right)^x\right]$$

$$= \frac{1}{x!}\lim_{r\to\infty}\left[(1+P)^{-r}\left(\frac{rP}{1+P}\right)^x\right]$$

$$= \frac{\lambda^x}{x!}\lim_{r\to\infty}\left[\left(1+\frac{\lambda}{r}\right)\right]^{-r} \lim_{r\to\infty}\left(1+\frac{\lambda}{r}\right)^{-x} \qquad [\because rP = \lambda]$$

$$= \frac{\lambda^x}{x!}\cdot e^{-\lambda}\cdot 1 = \frac{e^{-\lambda}\lambda^x}{x!}$$

which is the probability function of the Poisson distribution with parameter 'λ'.

**7·4·4. Probability Generating Function of Negative Binomial Distribution.** Let $X$ be a random variable following negative binomial distribution, then

$$P_X(s) = E(s^X) = \sum_{x=0}^{\infty} s^x\, p(x)$$

$$= \sum_{x=0}^{\infty} \binom{-r}{x} p^r (-qs)^x \qquad \text{[Using 7·21 } a)]$$

$$= p^r (1 - qs)^{-r} = [p/(1 - qs)]^r \qquad \ldots(7{\cdot}24)$$

**Example 7·45.** *An item is produced in large numbers. The machine is known to produce 5% defectives. A quality control inspector is examining the items by taking them at random. What is the probability that at least 4 items are to be examined in order to get 2 defectives ?*

**Solution.** If 2 defectives are to be obtained then it can happen in 2 or more trials. The probability of success is 0·05 for every trial. It is a negative binomial situation and the required probability is

$$= P(X = 4) + P(X = 5) + \ldots$$

$$= \sum_{x=4}^{\infty} \binom{x-1}{2-1} (0{\cdot}05)^2 (0{\cdot}95)^{x-2}$$

$$= 1 - \sum_{2}^{3} \binom{x-1}{2-1} (0{\cdot}05)^2 (0{\cdot}95)^{x-2}$$

$$= 1 - [(0{\cdot}05)^2 + 2(0{\cdot}05)^2(0{\cdot}95)]$$

$$= 0{\cdot}995$$

**Example 7·46.** *If $X \sim B(n, p)$ and $Y$ has negative binomial distribution with parameters $r$* and *$p$, prove that*

$$F_X(r - 1) = 1 - F_Y(n - r)$$

**[Delhi Univ. Spl. Course (Statistics Hons.), 1987]**

**Solution.**

$$1 - F_Y(n - r) = 1 - P(Y \le n - r) = P(Y > n - r)$$

$$= \sum_{n-r+1}^{\infty} \binom{y+r-1}{r-1} p^r \cdot q^y ; \; [z = y - (n - r + 1)]$$

$$= p^r q^{n-r+1} \cdot \sum_{z=0}^{\infty} \binom{z+n}{r-1} q^z$$

$$= p^r q^{n-r+1} \sum_{z=0}^{\infty} \left\{ \sum_{k=0}^{r-1} \binom{n}{k} \binom{z}{r-1-k} \right\} q^z$$

$$\left(\because \sum_{k=0}^{r} \binom{a}{k}\binom{b}{r-k} = \binom{a+b}{r}\right)$$

$$= p^r q^{n-r+1} \sum_{k=0}^{r-1} \left\{ \binom{n}{k} \sum_{z=r-1-k}^{\infty} \binom{z}{r-1-k} q^z \right\}$$

[Changing the order of summation and noting that $\binom{n}{r} = 0 ;\ n < r$]

$$= p^r q^{n-r+1} \sum_{k=0}^{r-1} \left[ \binom{n}{k} \sum_{t=0}^{\infty} \binom{t+r-1-k}{r-1-k} q^{t+r-1-k} \right],$$

$$t = z - (r - 1 - k)$$

$$= p^r q^n \sum_{k=0}^{r-1} \left\{ \binom{n}{k} q^{-k} \cdot (1-q)^{-(r-k)} \right\}$$

$$= \sum_{k=0}^{r-1} \binom{n}{k} p^k \cdot q^{n-k}$$

$$= P[X \leq (r-1)] = F_X(r-1)$$

**Example 7·47. (Banach's Match-box Problem).** *A certain mathematician always carries two match boxes (initially containing N match sticks). Each time he wants a match-stick he selects a box at random, inevitably a moment comes when he finds a box empty. Show that the probability that there are exactly r match-sticks in one box when the other box is found empty is*

$$\binom{2N-r}{n} \times \left(\frac{1}{2}\right)^{2N-r}$$

**Solution.** Let the two match boxes be numbered 1 and 2. Let the choice of the Ist box be regarded as failure and that of second box be regarded as a success.

Since the mathematician selects the match box at random,

$$p = \text{Probability of selecting second match box} = 1/2$$

$$\Rightarrow \qquad q = 1 - p = 1/2$$

The second box will be found empty if it is selected for the $(N + 1)$st time. At this stage, the first box will contain exactly $r$ matches if $(N - r)$ matches have already been drawn from it. Hence the second box will be found empty at the stage when the first box contains exactly $r$ matches if and only if $(N - r)$ failures precede the $(N + 1)$st success. Thus in a total of $N + 1 + (N - r) = 2N - r + 1$, trials the last one must be success and out of the remaining $(2N - r)$ trials we should have $(N - r)$ failures and $N$ successes.

$\therefore$ Probability that second box is found empty when there are exactly $r$ matches in first box is

$$= \binom{2N-r}{N}\left(\frac{1}{2}\right)^{N}\left(\frac{1}{2}\right)^{N-r}\frac{1}{2} = \binom{2N-r}{N}\left(\frac{1}{2}\right)^{2N-r+1}$$

Similarly, the probability that first box is found empty, when the second box contains exactly $r$ matches is given by

$$\binom{2N-r}{N}\left(\frac{1}{2}\right)^{2N-r+1}$$

Hence the required probability that one match box is found empty when the other contains exactly $r$ matches is

$$2 \times \binom{2N-r}{N}\left(\frac{1}{2}\right)^{2N-r+1} = \binom{2N-r}{N}\left(\frac{1}{2}\right)^{2N-r}$$

**Remark.** The statement that 'he finds the box empty' implies that when he used the last match in this box, he did not throw it away, but instead put it back in his pocket. Thus there is a difference between 'the box is empty' and 'the box is found empty'.

The box becomes empty when the $N$th match was taken from it but it is found to be empty only when it is selected for the $(N + 1)$st time.

**Example 7·48.** *$X$ is a negative binomial variate with p.f.*

$$f(x) = \begin{cases} \binom{k+x-1}{x} q^x p^k, & x = 0, 1, 2, \ldots \\ 0 & , \text{otherwise} \end{cases}$$

*Show that the moment recurrence formula is*

$$\mu_{r+1} = q\left[\frac{d\mu_r}{dq} + \frac{rk}{p^2}\mu_{r-1}\right]$$

*State how the moments of negative binomial variate can be written from the corresponding formulas for binomial variate.*

**[Punjab Univ. B. Sc. (Maths Hons.) 1990]**

**Solution.** For Negative Binomial Distribution with parameter $k$ and $p$,

$$\text{Mean} = k \cdot q/p = \mu, \text{ (say)}.$$

$$\therefore \quad \mu_r = \sum_{x=0}^{\infty}(x-\mu)^r f(x)$$

$$= \sum_{x=0}^{\infty}\left[\left(x - \frac{kq}{p}\right)^r \cdot \binom{k+x-1}{x} q^x \cdot p^k\right]$$

Differentiating w.r.to $q$, we get

$$\frac{d\mu_r}{dq} = \sum_x\left[r\left(x - \frac{kq}{p}\right)^{r-1} \times \left\{\frac{d}{dq}\left(x - \frac{kq}{p}\right)\right\}\binom{k+x-1}{x} q^x p^k\right]$$

$$+ \sum_x\left[\left(x - \frac{kq}{p}\right)^r \binom{k+x-1}{x}\left\{x q^{x-1} p^k + q^x \cdot k p^{k-1} \cdot \frac{dp}{dq}\right\}\right]$$

But $\dfrac{dp}{dq} = \dfrac{d}{dq}(1-q) = -1$

and $\dfrac{d}{dq}\left[x - \dfrac{kq}{p}\right] = \dfrac{d}{dq}\left[x - k\left(\dfrac{1}{p} - 1\right)\right] = \dfrac{k}{p^2}\cdot\dfrac{dp}{dq} = -\dfrac{k}{p^2}$

$$\therefore \quad \frac{d\mu_r}{dq} = -\frac{rk}{p^2}\sum_x \left(x - \frac{kq}{p}\right)^{r-1}\cdot f(x)$$

$$+ \sum_x \left(x - \frac{kq}{p}\right)^r \binom{k+x-1}{x} q^{x-1}p^k\left(x - \frac{kq}{p}\right)$$

$$= -\frac{rk}{p^2}\mu_{r-1} + \frac{1}{q}\sum_x \left(x - \frac{kq}{p}\right)^{r+1}\cdot f(x)$$

$$= -\frac{rk}{p^2}\mu_{r-1} + \frac{1}{q}\cdot\mu_{r+1}$$

$$\Rightarrow \quad \mu_{r+1} = q\left[\frac{d\mu_r}{dq} + \frac{rk}{p^2}\mu_{r-1}\right];\ r = 1, 2, 3, \ldots$$

**7·4·5. Deduction of Moments of Negative Binomial Distribution From Those of Binomial Distribution.** If we write

$$p = 1/Q,\ q = P/Q \text{ such that } Q - P = 1,$$

then the m.g.f. of negative binomial variate $X$ is given by [c.f. § 7·4·1] :

$$M_X(t) = (Q - Pe^t)^{-k} \qquad \ldots(*)$$

This is analogous to the m.g.f. of binomial variate $Y$ with parameters $n$ and $p'$, *viz.*,

$$M_Y(t) = (q' + p'e^t)^n;\ q' = 1 - p' \qquad \ldots(**)$$

Comparing (*) and (**), we get

$$q' = Q,\ p' = -P \text{ and } n = -k \qquad \ldots(***)$$

Using the formulae for moments of binomial distribution, the moments of negative binomial distribution are given by

$$\text{Mean} = np' = (-k)(-P) = kP$$

$$\text{Variance} = np'q' = (-k)(-P)Q = kPQ$$

$$\mu_3 = np'q'(q' - p') = (-k)(-P)Q(Q+P) = kPQ(Q+P)$$

$$\mu_4 = np'q'[1 + 3p'q'(n-2)]$$

$$= (-k)(-P)Q[1 + 3(-P)Q(-k-2)]$$

$$= kPQ[1 + 3PQ(k+2)]$$

**Example 7·49.** *Prove that the recurrence formula for negative binomial distribution is :* $f(x+1; r, p) = \dfrac{x+r}{x+1}\, q\, f(x; r, p)$

**(Utkal Univ. M.A., 1990)**

**Solution.** We have

$$f(x\,;r,p) = \binom{x+r-1}{r-1} p^r q^x$$

$$f(x+1\,;r,p) = \binom{x+r}{r-1} p^r q^{x+1}$$

$$\therefore \quad \frac{f(x+1\,;r,p)}{f(x\,;r,p)} = \frac{(x+r)\,!\,(r-1)\,!\,x\,!}{(r-1)\,!\,(x+1)\,!\,(x+r-1)\,!}\,q = \frac{x+r}{x+1}\,q$$

$$\Rightarrow \quad f(x+1\,;r,p) = \frac{x+r}{x+1}\cdot q\cdot f(x\,;r,p)$$

This recurrence relation is useful for fitting of the negative binomial distribution to the given data as discussed in the following example.

**Example 7·50.** *Given the hypothetical distribution :*

| *No. of cells :* $(x)$ | *0* | *1* | *2* | *3* | *4* | *5* | *Total* |
|---|---|---|---|---|---|---|---|
| *Frequency :* $(f)$ | *213* | *128* | *37* | *18* | *3* | *1* | *400* |

*Fit a negative binomial distribution and calculate the expected frequencies.*

**Solution.** $\mu_1' = \text{Mean} = \dfrac{\Sigma fx}{\Sigma f} = \dfrac{473}{400} = \cdot 6825 = \dfrac{r.q}{p}$ ...(*)

$$\mu_2' = \frac{511}{400} = 1{\cdot}2775,$$

$$\mu_2 = \mu_2' - \mu_1'^2 = 1{\cdot}2775 - (\cdot 6825)^2 = 0{\cdot}8117$$

$\therefore$ Variance $= 0{\cdot}8117 = \dfrac{rq}{p^2}$ ...(**)

Solving equations (*) and (**), we get

$$p = \frac{0{\cdot}6825}{0{\cdot}8117} = 0{\cdot}8408, \quad q = 1 - p = 0{\cdot}1592$$

$$\therefore \quad r = \frac{p \times 0{\cdot}6825}{q} = \frac{0{\cdot}5738}{0{\cdot}1592} = 3{\cdot}60456$$

$$f_0 = p^r = (\cdot 8408)^{3\cdot 6045} = 5352$$

$$f_1 = \frac{r+0}{0+1}\,q f_0 = rq f_0 = 0{\cdot}5738 \times 0{\cdot}5352 = 0{\cdot}3071$$

$$f_2 = \frac{r+1}{1+1}\cdot q\cdot f_1 = \frac{4{\cdot}60456}{2} \times 0{\cdot}1592 \times 0{\cdot}3071 = 0{\cdot}1126$$

$$f_3 = \frac{r+2}{2+1}\cdot q\cdot f_2 = \frac{5{\cdot}60456}{3} \times 0{\cdot}1592 \times 0{\cdot}1126 = 0{\cdot}0335$$

$$f_4 = \frac{r+3}{3+1}\cdot q\cdot f_3 = \frac{6{\cdot}60456}{4} \times 0{\cdot}1592 \times 0{\cdot}0335 = 0{\cdot}0088$$

$$f_5 = \frac{r+4}{4+1}\cdot q\cdot f_4 = \frac{7{\cdot}60456}{5} \times 0{\cdot}1592 \times 0{\cdot}0088 = 0{\cdot}000213$$

∴ Expected frequencies are :

| $Nf_0$ | $Nf_1$ | $Nf_2$ | $Nf_3$ | $Nf_4$ | $Nf_5$ |
|---|---|---|---|---|---|
| 214·0992 | 122·8596 | 45·0308 | 13·3928 | 3·5204 | 0·8524 |

| | | | | | | |
|---|---|---|---|---|---|---|
| ∴ *Observed Frequency* : | 213 | 128 | 37 | 18 | 3 | 1 |
| *Expected Frequency* : | 214 | 123 | 45 | 13 | 4 | 1 |

## EXERCISE 7 (c)

1. (*a*) Define negative binomial distribution. Give an example in which it occurs. Obtain its moment generating function. Hence or otherwise obtain its mean, variance and third central moment. [Gujarat Univ. B.Sc. 1992]

(*b*) If $X$ denotes the number of failures preceding the $r$th success in an infinite series of independent trials with constant probability $p$ of success for each trial, then identify the distribution of $X$ and obtain $E(X)$. What is the distribution when $r = 1$ ? [Delhi Univ. B. Sc. (Stat. Hons.), 1985]

2. (*a*). A well known baseball player has a lifetime batting average of 0·3. He needs 32 more hits to make up his lifetime total to 3000. What is the probability that 100 or fewer times at bat are required for him to achieve his goal?

(*b*) A scientist needs three diseased rabbits for an experiment. He has 20 rabbits availlable and inoculates them one at a time with a serum, quitting if and when he gets 3 positive reactions. If the probability is 0·25 that a rabbit can contract the disease from the serum, what is the probability that the scientist is able to get 3 diseased rabbits from 20 ?

3. A student has taken a 5 answer multiple choice examination orally. He continues to answer questions until he gets five correct answers. What is the probability that he gets them on or before the twenty-fifth question if he guesses at each answer ?

4. If a boy is throwing stones at a target, what is the probability that his 10th throw is his 5th hit, if the probability of hitting the target at any trial is 0·5 ?

5. In a series of independent trials with constant probability $p$ of success in each trial, show that the number of successes in a fixed number $n$ of independent trials follows a binomial distribution. Show further that the number of the trials required for a specified number $r$ of successes follows a negative binomial distribution. Obtain the mean and the variance of this distribution.

6. (*a*) Obtain the Poisson-distribution as a limiting case of the negative binomial distribution. [(Delhi Univ. B.Sc. (Stat Hons.) 1988]

(*b*) Show how the moments of negative binomial variate can be written from the corresponding formulae for the bibomial variate.

[Delhi Univ. B.Sc. (Maths Hons.), 1991]

7. Consider a sequence of Bernoulli trials with constant probability $p$ of success in a single trial. Find $P(x, k)$, the probability that exactly $x + k$ trials are required to get $k$ successes, $x = 0, 1, 2, \ldots$ Show that $P(x, k)$ defines the probability function of the discrete random variable $X$. Find the moment generating function of $X$. Hence find $E(X)$ and $V(X)$.

8. ( *a* ) Derive moment generating function of negative binomial distribution and hence show that mean < variance

( *b* ) Derive negative binomial distribution in the following form :

$$f(x) = \binom{-k}{x}(-P)^x Q^{-k-x}\ ;\ \begin{matrix} x = 0, 1, 2, \ldots \\ Q = 1 + P \end{matrix}$$

Obtain (*i*) moment generating function, mean and variance of this distribution.

(*ii*) Coefficient of skewness $\beta_1$.

(*iii*) Give an example of its occurrence. [Gujarat Univ. B.Sc. Oct. 1990)

9. Obtain the characteristic function of the negative binomial distribution given in the form:

$$f(x; \alpha, \lambda) = \binom{-\lambda}{x}\left(\frac{\alpha}{1+\alpha}\right)^{\lambda}\left(\frac{-1}{1+\alpha}\right)^{x}; x = 0, 1, 2, \ldots$$

and hence evaluate its first two moments.

10. ( *a* ) Show that for the negative binomial distribution $(Q - P)^{-r}$, where $Q - P = 1$, cumulant generating function $K(t) = -r \log[1 - P(e^t - 1)]$. Hence deduce that $\kappa_1 = rP$, $\kappa_2 = rPQ$ Also obtain $\kappa_3$ and $\kappa_4$.

[Delhi Univ. B.Sc. (Stat. Hons.) 1986]

( *b* ) Show that the mean deviation about mean for the negative binomial distribution is

$$2(\mu + 1)\binom{n+\mu}{\mu+1} p^{\mu+1} q^{-(n+\mu)}$$

where $\mu$ is the greatest integer contained in $np + 1$.

11. The number of accidents among 414 machine operators was investigated for three successive months. The following table gives the distribution of the operators according to the number $k$, of accidents which happened to the same operator. Fit the distribution of the type

$$P(X = k) = (-1)^k \binom{-v}{k} p^v q^k\ ; k = 0, 1, 2, \ldots, v > 0, q = 1 - p, 0 < p < 1$$

| $k$ | ... | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|---|
| Observed frequency | ... | 296 | 74 | 26 | 8 | 4 | 4 | 1 | 0 | 1 |

12. If $X$ has negative binomial distribution with parameters $(n, P)$, prove that $M_X(t) = (Q - Pe^t)^{-n}$. Hence find m.g.f. of $Z = (X - nP)/\sqrt{nPQ}$ and deduce that $Z$ is asymptotically normal as $n \to \infty$

**Hint.** Prove that $M_Z(t) \to \exp(t^2/2)$ as $n \to \infty$ [c.f. Example 7·19].

13. Let $Y$ have the negative binomial distribution. Let $X_j$ be the number of failures between the $(j - 1)$th and $j$th success. Then $\sum_{j=1}^{r} X_j = Y$. Find $E(Y)$, $\sigma_Y^2$ by obtaining the means and variances of the $X_j$'s.

**14.** Assume that the mutually independent random variables $X_i$, each have the negative binomial distribution with parameters $r_i$ ( $i = 1, 2, ..., n$ ), where $r_i$ are all positive integers, i.e.,

$$P(X_i = x) = \binom{r_i + x - 1}{x} p^{r_i} q^x \,;\; x = 0, 1, 2, \ldots$$

Then show that the probability density function of $\sum_{i=1}^{n} X_i$ is the negative binomial distribution with $r = \sum_{i=1}^{n} r_i$ *i.e.,* the negative binomial distribution (with fixed $p$) is reproductive with respect to $r$. **(Sagar Univ. M.A., 1991)**

**15.** Suppose that a radio tube is inserted into a socket and tested. Assume that the probability that it tests positive equals $P$ and the probability that it tests negative is $(1 - P)$. Assume furthermore that we are testing large supply of such tubes. The testing continues until the first positive tube appears. If $X$ is the number of tests required to terminate the experiment, what is the probability distribution of $X$ ? **[(Aligarh U. B.Sc. (Hons.) 1993)]**

**16.** A man buys two boxes of matches, each containing $N$ matches initially and places one match box in his right pocket and one in his left pocket. Every time when he wants a match, he selects a pocket at random. Show that the probability that at the moment when the first box is emptied (not found empty), the other box contain exactly $r$ matches ( $r = 1, 2, ..., N$ ) is

$$\binom{2N - 1 - r}{r - 1} \left(\frac{1}{2}\right)^{2N - r - 1}$$

Using this result, show that the probability that the box first emptied is not the one first found to be empty is

$$\left(\frac{1}{2}\right)^{2N} \cdot \sum_{r=1}^{N} \binom{2N - 1 - r}{N - 1},$$

which reduces to $\binom{2N}{N} \left(\frac{1}{2}\right)^{2N + 1}$ or $\frac{1}{2} (N\pi)^{-1/2}$ approximately.

**7·5. Geometric Distribution.** Suppose we have a series of independent trials or repetitions and on each repetition or trial the probability of success '$p$' remains the same. Then the probability that there are $x$ failures preceding the first success is given by $q^x p$.

**Definition.** *A random variable X is said to have a geometric distribution if it assumes only non-negative values and its probability mass function is given by*

$$P(X = x) = \begin{cases} q^x p \,;\; x = 0, 1, 2, \ldots, 0 < p \le 1 \\ 0, \text{ otherwise} \end{cases} \qquad \ldots(7{\cdot}25)$$

**Remarks. 1.** Since the various probabilities for $x = 0, 1, 2, \ldots,$ are the various terms of geometric progression, hence the name geometric distribution.

2. Clearly, assignment of probabilities in (7·25) is permissible, since

$$\sum_{x=0}^{\infty} P(X=x) = \sum_{x=0}^{\infty} q^x p = p(1+q+q^2+\ldots) = \frac{p}{1-q} = 1$$

**7·5·1. Lack of Memory.** The geometric distribution is said to lack memory in a certain sense. Suppose an event $E$ can occur at one of the times $t = 0, 1, 2, \ldots$ and the occurrence (waiting) time $X$ has a geometric distribution. Thus $P(X=t) = q^t \cdot p;\ t = 0, 1, 2, \ldots$

Suppose we know that the event $E$ has not occurred before $k$, *i.e.*, $X \geq k$. Let $Y = X - k$. Thus $Y$ is the amount of additional time needed for $E$ to occur. We can show that

$$P(Y=t \mid X \geq k) = P(X=t) = pq^t \qquad \ldots(7\cdot26)$$

which implies that the additional time to wait has the same distribution as initial time to wait.

Since the distribution does not depend upon $k$, it, in a sense, '*lacks memory*' of how much we shifted the time origin. If '$B$' were waiting for the event $E$ and is relieved by '$C$' immediately before time $k$, then the waiting time distribution of '$C$' is the same as that of '$B$'.

**Proof of (7·26).** We have

$$P(X \geq r) = \sum_{s=r}^{\infty} pq^s = p(q^r + q^{r+1} + \ldots) = \frac{pq^r}{(1-q)} = q^r$$

$$P(Y \geq t \mid X \geq k) = \frac{P(Y \geq t \cap X \geq k)}{P(X \geq k)} = \frac{P(X-k \geq t \cap X \geq k)}{P(X \geq k)}$$

$$(\because Y = X - k)$$

$$= \frac{P(X \geq k+t)}{P(X \geq k)} = \frac{q^{t+k}}{q^k} = q^t$$

$$\therefore P(Y=t \mid X \geq k) = P(Y \geq t \mid X \geq k) - P(Y \geq t+1 \mid X \geq k)$$

$$= q^t - q^{t+1} = q^t(1-q) = pq^t = P(X=t)$$

**7·5·2. Moments of Geometric Distribution.**

$$\mu_1' = \sum_{x=1}^{\infty} x \cdot P(X=x) = \sum_{x=1}^{\infty} x \cdot pq^x = pq \sum_{x=1}^{\infty} xq^{x-1} = pq(1-q)^{-2} = \frac{q}{p}$$

$$V(X) = E(X^2) - [E(X)]^2 = E[X(X-1)] + E(X) - [E(X)]^2$$

$$E[(X-1)X] = \sum_{x=1}^{\infty} x(x-1)P(X=x) = \sum_{x=2}^{\infty} x(x-1)pq^x$$

$$= 2pq^2 \sum_{x=2}^{\infty} \left[\frac{x(x-1)}{2 \times 1} q^{x-2}\right] = 2pq^2(1-q)^{-3} = \frac{2q^2}{p^2}$$

$$\therefore \quad V(X) = \mu_2 = \frac{2q^2}{p^2} + \frac{q}{p} - \frac{q^2}{p^2} = \frac{q^2}{p^2} + \frac{q}{p} = \frac{q}{p^2}$$

**7·5·3. Moment Generating Function of Geometric Distribution.**

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \cdot q^x p = p \sum_{x=0}^{\infty} (e^t q)^x = p(1-qe^t)^{-1}$$

$$= p/(1-qe^t) \quad \ldots(7\cdot27)$$

$$\mu_1' = \left[\frac{d}{dt} M(t)\right]_{t=0} = \left[\frac{d}{dt} p(1-qe^t)^{-1}\right]_{t=0}$$

$$= p\left[qe^t(1-qe^t)^{-2}\right]_{t=0} = pq(1-q)^{-2} = \frac{q}{p}$$

$$\mu_2' = \left[\frac{d^2}{dt^2} M(t)\right]_{t=0} = \frac{q}{p} + \frac{2q^2}{p^2} \quad \text{(On simplification)}$$

$$\mu_2 = \mu_2' - \mu_1'^2 = \frac{q}{p} + \frac{2q^2}{p^2} - \frac{q^2}{p^2} = \frac{q^2+pq}{p^2} = \frac{q}{p^2}$$

Hence the mean and variance of the geometric distribution are $q/p$ and $q/p^2$ respectively.

**Remark**. The p.g.f. of the geometric distribution is obtained on replacing $e^t$ by $s$ in (7·27) and is given by :

$$P_X(s) = p/(1-qs) \quad \ldots(7\cdot27\ a)$$

**Example 7·51.** *Let the two independent random variables $X_1$ and $X_2$ have the same geometric distribution. Show that the conditional distribution of $X_1 \mid (X_1 + X_2 = n)$ is uniform.*

**[Gujarat Univ. B.Sc. 1992; Calicut U. B.Sc. (Main Stat), Oct. 1990]**

**Solution.** We are given

$$P(X_1 = k) = P(X_2 = k) = pq^k ; k = 0, 1, 2 \ldots$$

$$P[X_1 = r \mid (X_1 + X_2 = n)] = \frac{P(X_1 = r \cap X_1 + X_2 = n)}{P(X_1 + X_2 = n)}$$

$$= \frac{P(X_1 = r \cap X_2 = n-r)}{P(X_1 + X_2 = n)}$$

$$= \frac{P(X_1 = r \cap X_2 = n-r)}{\sum_{s=0}^{n} [P(X_1 = s) \cap X_2 = n-s)}$$

$$= \frac{P(X_1 = r) \cdot P(X_2 = n-r)}{\sum_{s=0}^{n} [P(X_1 = s) \cdot P(X_2 = n-s)]}$$

[Since $X_1$ and $X_2$ are independent)

$$\therefore\ P[X_1 = r \mid (X_1 + X_2 = n)] = \frac{pq^r \cdot pq^{n-r}}{\sum_{s=0}^{n} [pq^s \cdot pq^{n-s}]} = \frac{p^2 q^n}{\sum_{s=0}^{n} (p^2 q^n)}$$

$$= \frac{p^2 q^n}{(n+1) p^2 q^n} = \frac{1}{n+1}; \; r = 0, 1, 2, \ldots n$$

Hence the conditional distribution of $X_1 \mid (X_1 + X_2 = n)$ is discrete uniform. (*cf.* § 7·8).

**Example 7·52.** *Suppose X is a non-negative integral valued random variable. Show that the distribution of X is geometric if it 'lacks memory', i.e., if for each $k \geq 0$ and $Y = X - k$ one has*

$$P(Y = t \mid X \geq k) = P(X = t), \text{ for } t \geq 0$$

[Madras Univ. B.Sc. (Main), 1988]

**Solution.** Let us suppose

$$P(X = r) = p_r; \; r = 0, 1, 2, \ldots$$

Define

$$q_k = P(X \geq k) = p_k + p_{k+1} + \ldots \qquad \ldots(*)$$

We are given

$$P(Y = t \mid X \geq k) = P(X = t) = p_t \qquad \ldots(**)$$

We have

$$P(Y = t \mid X \geq k) = \frac{P(Y = t \cap X \geq k)}{P(X \geq k)} = \frac{P(X - k = t \cap X \geq k)}{P(X \geq k)}$$

$$= \frac{P(X = k + t)}{P(X \geq k)} = \frac{p_{k+t}}{q_k}$$

$$\Rightarrow \qquad p_t = \frac{p_{k+t}}{q_k}, \qquad \text{[Using (**)]}$$

for every $t \geq 0$ and all $k \geq 0$. In particular, taking $k = 1$, we get

$$p_{t+1} = q_1 \cdot p_t = (p_1 + p_2 + \ldots) p_t = (1 - p_0) p_t \qquad \text{[From (*)]}$$

$$\Rightarrow \qquad p_t = (1 - p_0) p_{t-1} = (1 - p_0)^2 p_{t-2} = \ldots = (1 - p_0)^t p_0$$

Hence $\quad p_t = P(X = t) = p_0 (1 - p_0)^t; \; t = 0, 1, 2, \ldots$

$\Rightarrow$ $X$ has a geometric distribution.

## EXERCISE 7 (d)

1. (*a*) If the probability that a target is destroyed on any one shot is 0·5, what is the probaility that it would be destroyed on 6th attempt ?

Ans. $(0 \cdot 5)^6$

(*b*) A couple decides to have children until they have a male child. What is the probability distribution of the number of children they would have ? If the probability of a male child in their community is 1/3, how many children are they expected to have before the first male child is born ?

(Sardar Patel U.B.Sc. Nov. 1991)

(*c*) Let $X$ be a discrete random variable having geometric distribution with parameter $p$. Obtain its mean and variance. Also, show that for any two positive integers $s$ and $t$,

$$P[X > s + t | X > s] = P[X > t]$$

**2.** The following distribution relates to the number of accidents to 650 women working on highly explosive shells during 5-week period. Show that a negative binomial distribution, rather than a geometric distribution, gives a very good fit to the data. How would you explaint this ?

| Number of accidents : | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Frequency : | 450 | 132 | 41 | 22 | 3 | 2 |

(South Gujarat Univ. B.Sc. 1991)

**3.** (*a*) Show that the mean and variance of the geometric distribution

$$p(x) = q^x p \; ; \; x = 0\,1, 2, \ldots$$

are respectively $qp^{-1}$, $qp^{-2}$ (Allahabad Univ. B.Sc., 1989)

(*b*) Show that the mode of the distribution.

$$p(x) = (\tfrac{1}{2})^x ; x = 1, 2, 3, \ldots$$

is 1.

**4.** Find (*i*) the probability generating function, (*ii*) the moment generating function, and (*iii*) the cumulant generating function for discrete random variable $X$ following the geometric distribution

$$P(X = r) = (1 - p)\, p^{r-1} ; \; r = 1, 2, \ldots$$

**5.** $X_1$ and $X_2$ are independent random variables with the same distribution $q^k p \,; k = 0, 1 \ldots$ Let $Y$ be defined as the largest of $X_1$ and $X_2$, *i.e.*, $Y = \max(X_1, X_2)$. Obtain the joint distribution of $Y$ and $X_1$ and the distribution of $Y$.

**6.** Identify the distributions with the following M.G.F.

$$e^t (5 - 4e^t)^{-1}.$$

**Ans.** Geometric Distribution, $p = 1/5$,

**7.** Prove the recurrence formula for Geometric Distribution, *viz.*,

$$p(x + 1) = q.p(x)$$

Let $X$ and $Y$ be independent random variables such that

$$P(X = r) = P(Y = r) = q^r p \; ; \; r = 0, 1, 2, \ldots$$

$p$ and $q$ are positive numbers such that $p + q = 1$. Find (*i*) the distribution of $X + Y$ and (*ii*) the conditional distribution of $X$ given $X + Y = 3$.

**9.** A die is cast until 6 appears. What is the probability that it must be cast more than five times.

**Ans.** $P(X > 5) = 1 - P(X \le 5) = 1 - \sum_{x=1}^{5} (5/6)^{x-1} \cdot (1/6)$

**10.** For the geometric distribution with p.m.f.

$$f(x) = 2^{-x}\ ;\ x = 1, 2, 3, \ldots$$

show that Chebychev's inequality gives

$$P(|X - 2| \leq 2) > \tfrac{1}{2}$$

while the actual probability is 15/16.

**[Rajasthan Univ. B.Sc. (Hons.) 1992]**

**11.** The conditional distribution of random variable $X$ given $Y = y$ is $\dfrac{e^{-y} y^x}{x!}$ and the marginal probability density of $Y$ is $e^{-y}$, where $X$ is a discrete variable, *i.e.*, $x = 0, 1, 2, \ldots$ and $Y$ is continuous, $y \geq 0$.

Show that the marginal distribution of $X$ is geometric.

**Hint.** $g(x, y) = f(x \mid y)\ h(y) = \dfrac{e^{-y} y^x}{x!} \cdot e^{-y}$

$$\therefore \qquad f(x) = \int_0^{\infty} \frac{e^{-2y} y^x}{x!}\, dy = \frac{1}{x!} \int_0^{\infty} e^{-2y} y^x\, dy = \frac{1}{x!} \cdot \frac{x!}{2^{x+1}}$$

**12.** If $X$ and $Y$ be two independent random variables, each representing the number of failures preceding the first success in a sequence of Bernoulli trials with $p$ as probability of success in a single trial and $q$ as probability of failure, show that $P(X = Y) = \dfrac{p}{1 + q}$

**[Delhi Univ. B.Sc. (Stat. Hons.) 1993, '87]**

**Hint.** We have $P(X = r) = P(Y = r) = q^r \cdot p\ ;\ (r = 0, 1, 2, \ldots)$

$$P(X = Y) = \sum_{r=0}^{\infty} P(X = r \cap Y = r) = \sum_{r=0}^{\infty} P(X = r) \cdot P(Y = r)$$

[$\because$ $X$ and $Y$ are independent *r.v.'s* ]

$$= p^2 \sum_{r=0}^{\infty} q^{2r} = p^2(1 + q^2 + q^4 + \ldots) = \frac{p^2}{1 - q^2} = \frac{p}{1 + q}.$$

**7·6. Hypergeometric Distribution.** When the population is finite and the sampling is done without replacement, so that the events are stochastically dependent, although random, we obtain hypergeometric distribution. Consider an urn with $N$ balls, $M$ of which are white and $N - M$ are red. Suppose that we draw a sample of $n$ balls at random (*without replacement*) from the urn, then the probability of getting $k$ white balls out of $n$, $(k < n)$ is

$$\binom{M}{k}\binom{N - M}{n - k} \div \binom{N}{n}.$$

**Definition.** *A discrete random variable $X$ is said to follow the hypergeometric distribution if it assumes only non-negative values and its probability mass function is given by*

$$P(X=k)=h(k;N,M,n)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}};\ k=0,1,2,\ldots,\min(n,M).$$
$$=0,\ \text{otherwise} \qquad \ldots(7\cdot28).$$

**Remarks. 1.** $N$, $M$ and $n$ are known as the three parameters of hypergeometric distribution.

**2.** As it can be shown that

$$\sum_{k=0}^{n}\binom{M}{k}\binom{N-M}{n-k}\div\binom{N}{n}=1\ ,$$

this assignment of probabilities is permissible.

**7·6·1. Mean and Varaince of the Hypergeometric Distribution.**

$$E(X)=\sum_{k=0}^{n}k\cdot P(X=k)=\sum_{k=0}^{n}k\left\{\binom{M}{k}\binom{N-M}{n-k}\div\binom{N}{n}\right\}$$
$$=\frac{M}{\binom{N}{n}}\sum_{k=1}^{n}\left\{\binom{M-1}{k-1}\binom{N-M}{n-k}\right\}$$
$$=\frac{M}{\binom{N}{n}}\sum_{x=0}^{m}\binom{A}{x}\binom{N-A-1}{m-x},$$
$$\text{where } x=k-1,\ m=n-1,\ M-1=A$$
$$=\frac{M}{\binom{N}{n}}\cdot\binom{N-1}{m}=\frac{M}{\binom{N}{n}}\binom{N-1}{n-1}=\frac{nM}{N}$$

$$E\{X(X-1)\}=\sum_{k=0}^{n}k(k-1)\left\{\binom{M}{k}\binom{N-M}{n-k}\div\binom{N}{n}\right\}$$
$$=\frac{M(M-1)}{\binom{N}{n}}\sum_{k=2}^{n}\left\{\binom{M-2}{k-2}\binom{N-M}{n-k}\right\}$$
$$=\frac{M(M-1)}{\binom{N}{n}}\cdot\binom{N-2}{n-2}=\frac{M(M-1)n(n-1)}{N(N-1)}$$

---

*Since $k$ white balls can be drawn from '$M$' white balls in $\binom{M}{k}$ ways and out of the remaining $N-M$ red balls, $(n-k)$ can be chosen in $\binom{N-M}{n-k}$ ways, the total number of favourable cases is $\binom{M}{k}\times\binom{N-M}{n-k}$.

$$\therefore \quad E(X^2) = E[X(X-1)] + E(X) = \frac{M(M-1)n(n-1)}{N(N-1)} + \frac{nM}{N}$$

Hence $$V(X) = \frac{M(M-1)n(n-1)}{N(N-1)} + \frac{nM}{N} - \left(\frac{nM}{N}\right)^2$$

$$= \frac{NM(N-M)(N-n)}{N^2(N-1)} \quad \text{(On simplification)}$$

**7·6·2. Factorial Moments of Hypergeometric Distribution.** The $r$th factorial moment is

$$E[X^{(r)}] = \sum_{k=r}^{n} k^{(r)} P(X=k) = \sum_{k=r}^{n} k^{(r)} \left\{\binom{M}{k}\binom{N-M}{n-k} \div \binom{N}{n}\right\}$$

$$= \sum_{k=r}^{n} M^{(r)} \left\{\binom{M-r}{k-r}\binom{N-M}{n-k} \div \binom{N}{n}\right\}^{*}$$

$$= M^{(r)} \sum_{j=0}^{n-r} \left\{\binom{M-r}{j}\binom{(N-r)-(M-r)}{(n-r)-j} \div \binom{N}{n}\right\}, \text{ where } j = k-r$$

$$= \frac{M^{(r)} n^{(r)}}{N^{(r)}} \sum_{j=0}^{n-r} \left\{\binom{M-r}{j}\binom{(N-r)-(M-r)}{(n-r)-j} \div \binom{N-r}{n-r}^{**}\right\}$$

$$= \frac{M^{(r)} n^{(r)}}{N^{(r)}} \sum_{j=0}^{n-r} h(j; N-r, M-r, n-r) = \frac{M^{(r)} n^{(r)}}{N^{(r)}} \cdot 1$$

$$\therefore \quad E[X^{(r)}] = \frac{M^{(r)} n^{(r)}}{N^{(r)}} \qquad \ldots(7\cdot28\,a)$$

$$\Rightarrow \quad \mu_x = E(X) = \frac{nM}{N}$$

$$E[X^{(2)}] = \frac{M(M-1)n(n-1)}{N(N-1)}$$

$$\sigma_X^2 = E[X^{(2)}] + E(X) - [E(X)]^2 = n \cdot \frac{M}{N} \cdot \frac{N-M}{N} \cdot \frac{N-n}{N-1}$$

(On simplification) ...(7·28 b)

**Remark.** If we sample the $n$ balls with replacement and denote by $Y$ the number of white balls in the sample, then $Y$ is a binomial variate with parameters $n$ and $p$ where

---

* $\because k^{(r)}\binom{M}{k} = \frac{k(k-1)(k-2)\ldots(k-r+1)M!}{k!(M-k)!}$

$= \frac{M(M-1)(M-2)\ldots(M-r+1)(M-r)!}{(k-r)!(M-k)!} = M^{(r)}\binom{M-r}{k-r}$

** $n^{(r)}\binom{N}{n} = N^{(r)}\binom{N-r}{n-r}$

$$p = M/N,\ q = 1 - p = (N - M)/N$$

$$\therefore \qquad E(Y) = np = \frac{nM}{N} = E(X)$$

$$\sigma_Y^2 = npq = n \cdot \frac{M}{N} \cdot \frac{N-M}{N} \geq \sigma_X^2, \qquad \text{[From (7·28 } b\text{)]}$$

equality holding only if $n = 1$.

**7·6·3. Approximation to Binomial Distribution.** Hypergeometric distribution tends to binomial distribution as $N \to \infty$ and $\dfrac{M}{N} \to p$.

$$h(k; N, M, n) = \binom{M}{k}\binom{N-M}{n-k} \div \binom{N}{n}$$

$$= \frac{M!}{k!(M-k)!} \cdot \frac{(N-M)!}{(n-k)!(N-M-n+k)!} \cdot \frac{n!(N-n)!}{N!}$$

$$= \frac{M(M-1)(M-2)\dots(M-k+1)}{k!}$$

$$\times \frac{(N-M)(N-M-1)\dots(N-M-n+k+1)}{(n-k)!}$$

$$\times \frac{n!}{N(N-1)(N-2)\dots(N-n+1)}$$

$$= \frac{n!}{k!(n-k)!} \cdot \frac{M}{N}\left(\frac{M}{N} - \frac{1}{N}\right)\left(\frac{M}{N} - \frac{2}{N}\right)\dots\left(\frac{M}{N} - \frac{k-1}{N}\right)$$

$$\times \frac{\left(1 - \frac{M}{N}\right)\left(1 - \frac{M}{N} - \frac{1}{N}\right)\dots\left(1 - \frac{M}{N} - \frac{(n-k-1)}{N}\right)}{\left(1 - \frac{1}{N}\right)\left(1 - \frac{2}{N}\right)\dots\left(1 - \frac{n-1}{N}\right)}$$

Proceeding to the limit as $N \to \infty$ and putting $\dfrac{M}{N} = p$, we get

$$\lim_{N \to \infty} h(k; N, M, n) = \binom{n}{k} \underbrace{p \cdot p \dots p}_{k \text{ times}} \underbrace{(1-p)(1-p)\dots(1-p)}_{(n-k) \text{ times}}$$

$$= \binom{n}{k} p^k (1-p)^{n-k} = b(k; p, 1-p)$$

**7·6·4. Recurrence Relation for the Hypergeometric Distribution.** We have

$$h(k; N, M, n) = \binom{M}{k}\binom{N-M}{n-k} \div \binom{N}{n}$$

$$h(k+1; N, M, n) = \binom{M}{k+1}\binom{N-M}{n-k-1} \div \binom{N}{n}$$

$$\therefore \qquad \frac{h(k+1; N, M, n)}{h(k; N, M, n)} = \frac{(n-k)(M-k)}{(k+1)(N-M-n+k+1)},$$

which is the required recurrence relation.

**Example 7·53.** *Explain how you will use hypergeometric model to estimate the number of fish in a lake.*

**Solution.** Let us suppose that in a lake there are $N$ fish, $N$ unknown. The problem is to estimate $N$. A catch of '$r$' fish (all at the same time) is made and these fish are returned alive into the lake after marking each with a red spot. After a reasonable period of time, during which these marked' fish are assumed to have distributed themselves 'at random' in the lake, another catch of 's' fish (again, all at once) is made. Here $r$ and $s$ are regarded as fixed predetermined constants. Among these $s$ fish caught, there will be, (say), $X$ marked fish where $X$ is a random variable following discrete probability function given by hypergeometric model:

$$f_X(x \mid N) = \binom{r}{x}\binom{N-r}{s-r} \div \binom{N}{s} = p(N), \text{ say} \qquad \text{...(*)}$$

where $x$ is an integer such that $\max(0, s-N+r) \le x \le \min(r, s)$ and $f_X(x \mid N) = 0$ otherwise.

The value of $N$ is estimated by the principle of Maximum Likelihood (*c.f.* Chapter 15), *i.e.*, we find the value $\hat{N} = \hat{N}(x)$ of $N$ which maximises $p(N)$. Since $N$ is a discrete *r.v.*, the principle of maxima and minima in calculus cannot be used here. Here we proceed as follows :

$$\lambda(N) = \frac{p(N)}{p(N-1)} = \frac{(N-r)(N-s)}{N(N-r-s+x)}. \qquad \text{(On simplification)}$$

$$\therefore \ \lambda(N) > 1 \text{ iff } N > \frac{rs}{x} \Rightarrow p(N) > p(N-1) \text{ iff } N > \frac{rs}{x} \qquad \text{...}(i)$$

$$\text{and } \lambda(N) < 1 \text{ iff } N < \frac{rs}{x} \Rightarrow p(N) < p(N-1) \text{ iff } N < \frac{rs}{x} \qquad \text{...}(ii)$$

From (*i*) and (*ii*) we see that $p(N) = f_X(x \mid N)$ reaches the maximum value (as a function of $N$) when $N$ is approximately equal to $rs/x$. Hence maximum likelihood estimate of $N$ is given by

$$\hat{N} = \frac{rs}{x} \Rightarrow \hat{N}(X) = \frac{rs}{X}$$

## EXERCISE 7 (e)

1. (*a*) What is a hypergeometric distribution ? Find the mean and variance of this distribution. How is this distribution related to the binomial ?

[Nagarjuna Univ. M.Sc. 1991; Delhi Univ. B.Sc. (Stat. Hons.), 1989]

(*b*) Obtain binomial distribution as a limiting case of hyper-geometric distribution. [Delhi Univ. B.Sc. (Stat. Hons.), 1989,' 87]

2. Suppose that rockets of a certain type have, by many tests, been established as 90% reliable. Now a modification of the rocket design is being considered. Which of the following sets of evidence throws more doubt on the hypothesis that the modified rocket is only 90% reliable :

(*i*) Of 100 modified rockets tested, 96 performed satisfactorily.

(*ii*) Of 64 modified rockets tested, 62 performed satisfactorily ?

3. A taxi cab company has 12 Ambassadors and 8 Fiats. If 5 of these taxi cabs are in the shop for repairs and Ambassador is as likely to be in for repairs as a Fiat, what is the probbaility that ·

(i) 3 of them are Ambassadors and 2 are Fiats ?

(ii) at least 3 of them are Ambassadors ? and

(iii) all 5 of them are of the same make ?

Ans. (i) $\binom{12}{3}\binom{8}{2} \div \binom{20}{5}$; (ii) $\sum_{x=3}^{5}\binom{12}{x}\binom{8}{5-x} \div \binom{20}{5}$

4. (a) Show how the hypergeometric disrtribution arises, by giving an example. Obtain the frequency function of a random variable $X$ following the above law. Derive $E(X)$ and $V(X)$. Show that under certain conditions to be stated, the Binomial and Poisson distributions are special cases of the hypergeometric distribution. [Dibrugarh Univ. B.Sc. 1992]

(b) Find the factorial moments of the hypergeometric distribution.

[Delhi Univ. B.Sc. (Stat Hons.), 1993]

5. (a) Suppose that from a population of $N$ elements of which $M$ are defective and $(N-M)$ are non-defective, a sample of size $n$ is drawn without replacement. What is the probability that the sample contains exactly $x$ defectives? Name this probability distribution.

(b) Show that, for the distribution derived in (a),

$$E(X) = \frac{nM}{N} \text{ and } (ii)\ V(X) = \frac{nM}{N}\left(1 - \frac{M}{N}\right)\left(1 - \frac{n-1}{N-1}\right)$$

(c) Show that, under certain conditions to be stated, the binomial distribution may be looked upon as a limiting form of the probability distribution as derived in (a).

6. (a) 200 students of the F.Y. B.Sc. class in a certain College are divided at random into 20 batches of 10 each for the annual practical examination in Statistics. Suppose the class consists of 40 resident students and 160 non-resident students; and let R denote the number of resident students in the first batch. Use the binomial approximation to find the probability that $R \geq 3$.

**Hint.** The probability distribution of $R$ is hyper-geometric with parameters : $N = 200$, $n = 10$, $M = 40$

Since $N$ $(= 200)$ is large, the hypergeometric distribution (*) can be approximated by binomial distribution with parameters $n = 10$, $p = M/N = 40/200 = 0{\cdot}2$

$$\therefore \quad P(R = r) = \binom{10}{r}(0{\cdot}2)^r\,(0{\cdot}8)^{10-r};\ r = 0, 1, \ldots, 10,$$

and required probability is :

$$P(R \geq 3) = 1 - [P(R = 0) + P(R = 1) + P(R = 2)] = 0{\cdot}323$$

(b) Find the probability that the income-tax official will catch 3 income-tax returns with illegitimate deductions, if he randomly selects 5 returns from among 12 returns of which 6 contain illegitimate deductions.

**Ans.** $\binom{6}{3}\binom{6}{2} \div \binom{12}{5} = 25/66.$

(c) If $X$ and $Y$ are independent binomial variates with parameters $(n_1, p)$ and $(n_2, p)$ respectively, find $P(X = r \mid X + Y = n)$

**Ans.** $\binom{n_1}{r}\binom{n_2}{n-r} \div \binom{n_1+n_2}{n}$

**7.** From a finite population of $N$ animals in a given region, W are caught, marked and then released again. The animals are caught again one by one until m (pre-assigned) marked animals are caught. The total number of animals caught is a random variable $X$. Find P (X = n), for $m \le n \le N - W + m$

(Shivaji Univ. B.Sc., 1987)

**Hint.** $P(X = n) = P$ {Catching $(n-1)$ animals of whom $(m-1)$ are marked} $\times$ $P$ {Catching the marked animal from the remaining $N - (n-1)$ animals}.

$$= \frac{\binom{W}{m-1}\binom{N-W}{n-m}}{\binom{N}{n-1}} \times \frac{\{W - (m-1)\}}{\{N - (n-1)\}}$$

$$= \binom{N-n}{W-m}\binom{n-1}{m-1} \div \binom{N}{W}$$

**8.** An urn contains $M$ balls numbered 1 to $M$, where the first k balls are defective and the remaining $(M-K)$ are non-defective. A sample of $n$ balls is drawn from the urn. Let $A_k$ be the event that the sample of $n$ balls contains exactly $k$ defectives. Find $P(A_k)$ when the sample is drawn (i) with replacement and (ii) without replacement. [Delhi Univ. B.Sc. (Maths Hons.), 1989]

**Hint.** If sampling is done without replacement, we get hyper-geometric probability model.

$$P(A_k) = \binom{K}{k}\binom{M-K}{n-k} \div \binom{M}{n}$$

If sampling is done with replacement, then $X \sim B(n,\ p = K/M)$

$$\therefore\ P(A_k) = \binom{n}{k}(K/M)^k \cdot \left(1 - \frac{K}{M}\right)^{n-k} = \binom{n}{k}\frac{K^k (M-K)^{n-k}}{M^n}$$

**9.** $X$ is a random variable distributed according to hyper-geometric law:

$$P(X = x) = h(x; n, a, b) = \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}};\ x = 0, 1, 2, \ldots$$

Obtain the recurrence formula :

$$h(x+1; n, a, b) = \frac{(n-x)(a-x)}{(x+1)(b-n+x+1)}\, h(x; n, a, b)$$

**10.** For the hypergeometric distribution

$$h(N\,;n,p,x)=\frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}\ ;\ x=0,1,2,\dots$$

Prove that $\mu_1'=np$ and $\mu_2=\dfrac{n(N-n)pq}{N-1}$.

**11.** Explain how you will use hypergeometric model to estimate the number of wild animals in a dense forest.

**12.** A box contains $N$ items of which '$a$' items are defective and '$b$' are non-defective, $(a+b=N)$. A sample of $n$ items is drawn at random. Let $X$ be number of defective items in the sample. Obtain the probability distribution of $X$ and obtain the mean of the distribution.

**7·7. Multinomial Distribution.** This distribution c[illegible] regarded as a generalisation of Binomial distribution.

When there are more than two mutually exclusive out[illegible] trial, the observations lead to multinomial distribution. Suppose [illegible] $E_k$ are $k$ mutually exclusive and exhaustive outcomes of a trial with [illegible]ive probabilities $p_1, p_2, \dots, p_k$.

The probability that $E_1$ occurs $x_1$ times, $E_2$ occurs $x_2$ ti[illegible] ... and $E_k$, occurs $x_k$ times in $n$ independent observations, is given by

$$p(x_1,x_2,\dots,x_k)=c\,p_1^{x_1}\,p_2^{x_2}\dots p_k^{x_k}$$

where $\sum x_i=n$ and $c$ is the number of permutation of the events $E_1, E_2, \dots,$ [illegible]

To determine $c$, we have to find the number of permutations of $n$ objects [illegible] which $x_1$ are of one kind, $x_2$ of another kind, ..., and $x_k$ of the $k$th kind, which is given by

$$c=\frac{n\,!}{x_1\,!\,x_2\,!\dots x_k\,!}$$

Hence

$$p(x_1,x_2,\dots,x_k)=\frac{n!}{x_1\,!\,x_2\,!\dots x_k\,!}\cdot p_1^{x_1}\,p_2^{x_2}\dots p_k^{x_k}\ ,\ 0\le x_i\le n$$

$$=\frac{n\,!}{\prod\limits_{i=1}^{k}x_i\,!}\cdot\prod_{i=1}^{k}p_i^{x_i}\,,\qquad \sum_{i=1}^{k}x_i=n \qquad\dots(7\cdot29)$$

which is the required probability function of the multinomial distribution. It is so-called since (7·29) is the general term in the multinomial expansion

$$(p_1+p_2+\dots+p_k)^n,\ \sum_{i=1}^{k}p_i=1$$

Since, total probability is 1, we have

$$\sum_x p(x)=\sum_x\left[\frac{n!}{x_1\,!\,x_2\,!\dots x_k\,!}p_1^{x_1}\,p_2^{x_2}\dots p_k^{x_k}\right]=(p_1+p_2\dots+p_k)^n=1. \qquad\dots[7\cdot29(a)]$$

**7·7·1. Moments of Multinomial Distribution.** The moment generating function is given by

$$M_X(t) = M_{X_1, X_2, \ldots, X_k}(t_1, t_2, \ldots, t_k) = E\left[\exp\left\{\sum_{i=1}^{k} t_i X_i\right\}\right]$$

$$= \sum_{x}\left[\frac{n!}{x_1!\,x_2!\ldots x_k!}\,p_1^{x_1}\,p_2^{x_2}\ldots p_k^{x_k}\exp\left(\sum_{i=1}^{k} t_i x_i\right)\right]$$

$$= \sum_{x}\left[\frac{n!}{x_1!\,x_2!\ldots x_k!}\,(p_1 e^{t_1})^{x_1}\ldots(p_k e^{t_k})^{x_k}\right]$$

$$= (p_1 e^{t_1} + p_2 e^{t_2} + \ldots + p_k e^{t_k})^n \qquad \ldots(7\cdot30)$$

where $x = (x_1, x_2, \ldots, x_k)$ · [ On using (7·29 (a) ],

Now $M_{X_1}(t_1) = M_X(t_1, 0, 0, \ldots, 0) = \left(p_1 e^{t_1} + p_2 + p_3 + \ldots + p_k\right)^n$

$$= \left[(1 - p_1) + p_1 e^{t_1}\right]^n \qquad (\because \sum_i p_i = 1)$$

$\Rightarrow X_1 \sim B(n, p_1)$ [By uniqueness theorem of m.g.f.]

Similarly, we shall get:

$X_i \sim B(n, p_i)$ ; $i = 1, 2, \ldots, k$.

$\Rightarrow \; E(X_i) = np_i$ and $\text{Var}\, X_i = np_i(1 - p_i)$ ; $i = 1, 2, \ldots, k$

$$E(X_i X_j) = \left[\frac{\partial^2 M}{\partial t_i\,\partial t_j}\right]_{t=0}, \; i \neq j$$

$$= \left[np_i e^{t_i}\,(n-1)(p_1 e^{t_1} + \ldots + p_k e^{t_k})^{n-2}\,p_j e^{t_j}\right]_{t=0}$$

$$= n(n-1)p_i p_j$$

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = n(n-1)p_i p_j - n^2 p_i p_j = -np_i p_j$$

$$\therefore \rho(X_i, X_j) = \frac{\text{Cov}(X_i, X_j)}{\sigma_{X_i}\,\sigma_{X_j}} = \frac{-np_i p_j}{\sqrt{np_i(1-p_i)\,np_j(1-p_j)}}$$

$$= -\left[\frac{p_i p_j}{(1-p_i)(1-p_j)}\right]^{1/2}$$

**Example 7·54.** *The trinomial distribution of two r.v.'s X and Y is given by :*

$$f_{X,Y}(x, y) = \frac{n!}{x!\,y!\,(n-x-y)!}\,p^x q^y (1 - p - q)^{n-x-y}$$

*for* $x, y = 0, 1, 2, \ldots, n$ *and* $x + y \leq n$,

*where* $0 \leq p$, $0 \leq q$ and $p + q \leq 1$.

(*i*) *Find the marginal distributions of X* and *Y*

(*ii*) *Find the conditional distributions of X* and *Y and obtain* $E(Y \mid X = x)$ *and* $E(X \mid Y = y)$

(*iii*) *Find the correlation coefficient between X and Y.*

**[Delhi Univ. B.Sc. (Maths Hons.) 1988; Spl Course-Statistics 1989;' 85]**

**Solution.** The joint m.g.f. of $X$ and $Y$ is given by :

$$M_{X,Y}(t_1,t_2) = E(e^{t_1X+t_2Y}) = \sum_{x=0}^{n}\sum_{y=0}^{n-x}(pe^{t_1})^x(qe^{t_2})^y(1-p-q)^{n-x-y}$$

$$= \left[pe^{t_1} + qe^{t_2} + (1-p-q)\right]^n \qquad ...(i)$$

$$M_X(t_1) = M(t_1,0) = \{(1-p) + pe^{t_1}\}^n \Rightarrow X \sim B(n,p) \qquad ...(ii)$$

$$M_Y(t_2) = M(0,t_2) = \{(1-q) + qe^{t_2}\}^n \Rightarrow Y \sim B(n,q) \qquad ...(iii)$$

Observe that $M(t_1,t_2) \neq M(t_1,0) \times M(0,t_2) \Rightarrow X$ and $Y$ are not independent.

(ii) The conditional distribution of $X$ given $Y = y$ is given by :

$$f(X|Y=y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{f_{XY}(x,y)}{{}^nC_y\, q^y(1-q)^{n-y}} \qquad [\because Y \sim B(n,q)]$$

$$= \frac{(n-y)!}{x!\,(n-y-x)!}\left(\frac{p}{1-q}\right)^x\left(\frac{1-p-q}{1-q}\right)^{n-y-x}$$

$$= \binom{n-y}{x}\left(\frac{p}{1-q}\right)^x\left(1-\frac{p}{1-q}\right)^{n-y-x} \quad ; x = 0, 1, ..., n$$

$$\Rightarrow \quad X|(Y=y) \sim B(n-y,\ p/(1-q)) \qquad ...(iv)$$

$$\Rightarrow E(X|(Y=y)) = (n-y)\cdot p/(1-q) \qquad ...(v)$$

Similarly, we shall get

$$f(Y|X=x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{f(x,y)}{{}^nC_x\, p^x(1-p)^{n-x}} \qquad [\because X \sim B(n,p)]$$

$$= \binom{n-x}{y}\left(\frac{q}{1-p}\right)^y\left(1-\frac{q}{1-p}\right)^{n-x-y} \quad ; y = 0, 1, ..., n.$$

$$\Rightarrow \quad Y|(X=x) \sim B(n-x,\ q/(1-p)) \qquad ...(vi)$$

$$\Rightarrow E[Y|(X=x)] = (n-x)\, q/(1-p) \qquad ...(vii)$$

*(iii) Correlation Coefficient* $\rho_{XY}$ :

Since $X \sim B(n,p)$, $E(X) = np$, $Var\ X = np(1-p)$

$Y \sim B(n,q)$, $E(Y) = nq$, $Var\ Y = nq(1-q)$

$$E(XY) = \left.\frac{\partial^2 M(t_1,t_2)}{\partial t_1\,\partial t_2}\right|_{t_1=t_2=0} = n(n-1)pq$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = n(n-1)pq - n^2pq = -npq$$

$$\therefore\ \rho_{XY} = \frac{Cov\,(X, Y)}{\sigma_X\,\sigma_Y} = \frac{-npq}{\sqrt{np\,(1-p)\,nq\,(1-q)}} = -\left[\frac{pq}{(1-p)(1-q)}\right]^{1/2}$$

**Note.** Here $p + q \neq 1$.

**Example 7·55.** *If $X_1, X_2, ..., X_k$ are $k$ independent Poisson variates with prameters $\lambda_1, \lambda_2, ..., \lambda_k$ respectively, prove that the conditional distribution $P(X_1 \cap X_2 \cap ... \cap X_k \mid X)$, where $X = X_1 + X_2 + .... + X_k$ is fixed, is multinomial.* **[Lucknow U. B.Sc. (Hons.), 1992]**

**Solution.** $P[X_1 \cap X_2 \cap ... \cap X_k \mid X = n]$

$$= P[X_1 = r_1 \cap X_2 = r_2 \cap ... \cap X_k = r_k \mid X = n]$$

$$= \frac{P[X_1 = r_1 \cap X_2 = r_2 \cap ... \cap X_k = r_k \cap X = n]}{P(X = n)}$$

$$= \frac{P[X_1 = r_1 \cap ... \cap X_{k-1} = r_{k-1} \cap X_k = n - r_1 - r_2 ... - r_{k-1}]}{P(X = n)}$$

$$= \frac{P(X_1 = r_1)\,P(X_2 = r_2)...P(X_{k-1} = r_{k-1})\,P(X_k = n - r_1 - ... - r_{k-1})}{P(X = n)}$$

($\because X_1, X_2, ..., X_k$ are independent)

Further, since $X_i$, $(i = 1, 2, ..., k)$ are independent Poisson variates with parameters $\lambda_i$ respectively, $X = X_1 + X_2 + ... + X_k$ is also a Poisson variate with parameter $\lambda_1 + \lambda_2 + ... + \lambda_k = \lambda$ (say).

Hence $P[X_1 \cap X_2 \cap ... \cap X_k \mid X = n]$

$$= \frac{\dfrac{e^{-\lambda_1}\lambda_1^{r_1}}{r_1!} \cdots \dfrac{e^{-\lambda_{k-1}}\lambda_{k-1}^{r_{k-1}}}{r_{k-1}!} \cdot \dfrac{e^{-\lambda_k}\lambda_k^{n-r_1-...-r_{k-1}}}{(n-r_1-...-r_{k-1})!}}{\dfrac{e^{-\lambda}\lambda^n}{n!}}$$

$$= \left[\frac{n!}{r_1!\,r_2!\,...\,r_{k-1}!\,(n-r_1-...-r_{k-1})!}\right] \times \left[\left(\frac{\lambda_1}{\lambda}\right)^{r_1} \cdots \left(\frac{\lambda_{k-1}}{\lambda}\right)^{r_{k-1}} \left(\frac{\lambda_k}{\lambda}\right)^{n-r_1-...-r_{k-1}}\right]$$

$$= \frac{n!}{r_1!\,r_2!\,...\,r_k!}\,p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k}$$

where $\sum_{i=1}^{k} r_i = n$ and $\sum_{i=1}^{k} p_i = \sum_{i=1}^{k}\left(\frac{\lambda_i}{\lambda}\right) = \frac{1}{\lambda}\sum_{i=1}^{k}\lambda_i = 1$

Thus the conditional distribution $P(X_1 \cap X_2 \cap ... \cap X_k \mid X = n$ multinomial with probabilities $p_i = (\lambda_i/\lambda)$; $i = 1, 2, ..., k$ in $k$ classes.

**Remark.** If $X_i$'s are identically distributed independent Poisson variates with parameter $m$ (say), then $\lambda_i = m$; $i = 1, 2, ..., k$ and $\lambda = \sum_{i=1}^{k} \lambda_i = km$.

$$\therefore \qquad p_i = \frac{\lambda_i}{\lambda} = \frac{1}{k}$$

Hence in this case the conditional distribution of $X_1, X_2, ..., X_k$, given that their sum $X_1 + X_2 + ... + X_k = n$, is a multinomial distribution with index $n$ and the probability in each class being equal to $1/k$.

## EXERCISE 7(f)

1. If $X_1, X_2, ..., X_k$ have a multinomial distribution with the parameters $n$ and $p_i$ $(i = 1, 2, ..., k)$ with $\Sigma p_i = 1$, obtain the joint probability

$$P(X_1 = x_1 \cap X_2 = x_2 \cap ... \cap X_k = x_k)$$

Obtain the corresponding moment generating function. Hence, or otherwise show that $E(X_i) = np_i$, $V(X_i) = np_i(1 - p_i)$

and $$\text{Cov}(X_i, X_j) = -np_i p_j, \ (i \neq j).$$

2. Discuss the marginal and conditional distributions, associated with the multinomial distribution. If $(n_1, n_2, ..., n_k)$ have a multinomial distribution with parameters $(n, p_1, p_2, ..., p_k)$ and if $c_i, d_i, i = 1, 2, ... k$ are constants, find the variance of $\sum_{i=1}^{k} c_i n_i$ and co-variance between $\sum_{i=1}^{k} c_i n_i$ and $\sum_{i=1}^{k} d_i n_i$.

3. If the random variables $X_1, X_2, ..., X_k$ have a multinomial distribution, show that the marginal distribution of $X_i$ is a binomial distribution with the parameters $n$ and $p_i$, with $i = 1, 2, ..., k$.

4. For the trinomial distribution of two r.v.'s $X$ and $Y$ given by:

$$f(x, y) = \frac{n!}{x!\,y!\,(n - x - y)!} p_1^x p_2^y p_3^{n-x-y}$$

where $x$ and $y$ are non-negative integers with $x + y \leq n$ and $p_1, p_2, p_3$ are proper positive fractions with $p_1 + p_2 + p_3 = 1$, and

$f(x, y) = 0$, otherwise

Show that (*i*) $X \sim B(n, p_1)$ and $Y \sim B(n, p_2)$

(*ii*) $X | (Y = y) \sim B(n - y, p_1/(1 - p_2))$ and $Y | (X = x) \sim B(n - x, p_2/(1 - p_1))$

(*iii*) $\rho(X, Y) = -\left[p_1 p_2/(1 - p_1)(1 - p_2)\right]^{1/2}$

4. If '$n$' dice, each of which has 6 faces marked 1 to 6 are thrown, find the probability of getting a sum '$s$' on them.

**Hint.** The exhaustive number of ways in which $n$ dice can fall is $6^n$.

Since the total number of permutations in which six numbers, *viz.*, $1, 2, ..., 6$ taken '$n$' at a time can add to $s$ is the coefficient of $x^s$ in the multinomial expansion of $(x + x^2 + ... + x^6)^n$, the required number of favourable cases for getting a sum '$s$' on a dice is the co-efficient of $x^s$ in the expansion of $(x + x^2 + \ldots + x^6)^n$.

$\therefore$ Required Probability $= \frac{1}{6^n}$ [co-efficient of $x^s$ in $(x + x^2 + \ldots x^6)^n$ ]

Now identically, we have :

$$x + x^2 + \ldots + x^6 = x(1 + x + \ldots + x^5) = x\left(\frac{1 - x^6}{1 - x}\right)$$

and by binomial expansion

$$x^n (1 - x^6)^n = \sum_{k=0}^{n} (-1)^k . {}^nC_k \, x^{n+6k}$$

and
$$(1-x)^{-n} = \sum_{r=0}^{\infty} {}^{-n}C_r (-x)^r = \sum_{r=0}^{\infty} {}^{n+r-1}C_r . x^r = \sum_{r=0}^{\infty} {}^{n+r-1}C_{n-1} . x^r$$

$$\therefore \quad \left[x\left(\frac{1-x^6}{1-x}\right)\right]^n = \sum_{k=0}^{n} \sum_{r=0}^{\infty} (-1)^k . {}^nC_k . {}^{n+r-1}C_{n-1} . x^{n+6k+r}$$

To find the co-efficient of $x^s$, we put $n + 6k + r = s$ *i.e.*, $n + r = s - 6k$

Thus the co-efficient of $x^s$ in $(x + x^2 + \ldots + x^6)^n$

$$= \sum_{k=0}^{(s-n)/6} (-1)^k . {}^nC_k . {}^{s-6k-1}C_{n-1},$$

summation being extended over the integral values of $k$ not exceeding $(s - n)/6$.

Hence required probability $= \sum_{k=0}^{(s-n)/6} (-1)^k . {}^nC_k \, {}^{s-6k-1}C_r/6^n$

**Remarks. 1.** The probability of getting a sum '$s$' with a throw of $n$ dice, each having '$f$' faces marked 1 to $f$ is the co-efficient of $x^s$ in

$$\frac{1}{f^n}\left[\left(x + x^2 + \ldots + x^f\right)^n\right]$$

**2.** If $n$ dice have faces $f_1, f_2, \ldots, f_n$ respectively, then the required probability of getting a sum '$s$' is the co-efficient of $x^s$ in

$$\frac{1}{f_1, f_2 \ldots f_n}\left[(x + x^2 + \ldots + x^{f_1})(x + x^2 + \ldots + x^{f_2}) \ldots (x + x^2 + \ldots + x^{f_n})\right]$$

**6.** What is the probability of obtaining a sum of 15 points by throwing five dice toegther?

**Hint.** The number of exhaustive cases in throwing of 5 dice is $6^5$.

The number of ways in which the 5 dice thrown will give 15 points is the co-efficient of $x^{15}$ in the expansion of $(x^1 + x^2 + x^3 + \ldots + x^6)^5$.

Favourable number of cases

$\therefore$ $=$ coefficient of $x^{15}$ in $(x + x^2 + \ldots + x^6)^5$

$=$ coefficient of $x^{10}$ in $(1 + x + \ldots + x^5)^5$

$= \text{coefficient of } x^{10} \text{ in } (1-x^6)^5 (1-x)^{-5}$

$$(1-x^6)^5 = (1 - {}^5C_1 x^6 + {}^5C_2 x^{12} - \ldots - x^{30}) = (1 - 5x^6 + 10x^{12} - \ldots - x^{30})$$

$$(1-x)^{-5} = 1 + 5x + \frac{5 \times 6}{2\,!} x^2 + \frac{5 \times 6 \times 7}{3\,!} x^3 + \frac{5 \times 6 \times 7 \times 8}{4\,!} x^4 + \ldots$$

$$+ \frac{5 \times 6 \times 7 \times \ldots \times 14}{10\,!} x^{10} + \ldots$$

$$= (1 + 5x + 15x^2 + 35x^3 + 70x^4 + \ldots + 1001x^{10} + \ldots)$$

$\therefore$ Favourable number of cases

$= \text{co-efficient of } x^{10} \text{ in } (1 - 5x^6 + 10x^{12} - \ldots - x^{30})$

$\times (1 + 5x + \ldots + 70x^4 + \ldots + 1001x^{10} + \ldots)$

$= (1001 - 5 \times 70) = 651$

Hence the required probability $= \dfrac{651}{6^5} = \dfrac{651}{7776}$

7. Four dice, each marked 1 to 6, are thrown together. Find the probability of a total count being

(*i*) Exactly 12 and

or (*ii*) More than or equal to 20.

8. Four tickets marked 00,01,10, 11 respectively are placed in a bag. A ticket is drawn at random five times, being replaced each time. Find the probability that the sum of the numbers on the tickets thus drawn is 23.

9. Show that the mode of the multinomial distribution is given by $x_1, x_2, \ldots x_k$, satisfying

$$np_i - 1 < x_i \le (n + k - 1)\, p_i\,;\; i = 1, 2, \ldots, k$$

[In order to establish this, show that

$$p_i x_j \le p_j (x_i + 1) \text{ for } 1 \le i, j \le k]$$

**7·8. Discrete Uniform Distribution.** A random variable $X$ is said to have uniform distribution on $n$ points $\{x_1, x_2, \ldots, x_n\}$ if its p.m.f. is given by :

$$P(X = x_i) = \frac{1}{n}\,;\; i = 1, 2, \ldots, n \qquad \ldots(7\cdot31)$$

For example, if $X$ has a uniform distribution on the points $\{0, 1, 2, \ldots, n\}$, then $P(X = i) = \dfrac{1}{n+1}\,;\; i = 0, 1, 2, \ldots, n$. $\qquad \ldots(7\cdot31\,a)$

Such distributions can be conceived in practice if under the given experimental conditions, the different values of the random variable become euqally likely. Thus for a die experiment, and for an experiment with a deck of cards such distribution is appropriate.

**7·9. Power Series Distribution.** A discrete r.v. $X$ is said to follow a *generalized power series distribution* (g.p.s.d.), if its probability mass function is given by—

$$P(X=x)=\begin{cases}\dfrac{a_x\,\theta^x}{f(\theta)} & ;\ x=0,1,2,\ldots;\ a_x\ge 0\\[2ex] 0, \text{ elsewhere}\end{cases} \qquad \ldots(7\cdot32)$$

where $f(\theta)$ is a generating function, *i.e.*,

$$f(\theta)=\sum_{x\in S} a_x\,\theta^x,\ \theta\ge 0 \qquad \ldots(7\cdot32a)$$

so that $f(\theta)$ is positive, finite and differentiable and $S$ is a non-empty countable sub-set of non-negative integers.

**Remarks 1.** By taking proper choice of $S$ and $f(\theta)$, the g.p.s.d. can be reduced to binomial, Poisson and logarithmic series distribution and their truncated forms.

2. An inflated powerseries distribution (p.s.d.), inflated at zero is given by

$$P(X=x)=\begin{cases}1-\alpha+\dfrac{\alpha\,a_0}{f(\theta)}, & x=0\\[2ex] \alpha\,\dfrac{a_x\,\theta^x}{f(\theta)}; & x=1,2,\ldots\end{cases} \qquad \ldots(7\cdot33)$$

where $\alpha$ $(0<\alpha\le 1)$, is the inflation parameter.

3. The truncated p.s.d. is given by:

$$P(X=x\mid S)=\frac{a_x\,\theta^x}{f(\theta)}\Big/ f(S),\ x\in S$$
$$=0, \text{ otherwise}$$

$$\Rightarrow \quad P(X=x\mid S)=\frac{a_x\,\theta^x}{f_1(\theta)};\ x\in S, \text{ where } f_1(\theta)=\sum_{x\in S} a_x\,\theta^x$$
$$=0, \text{ otherwise} \qquad \ldots(7\cdot34)$$

**7·9·1. Moment Generating Function of p.s.d.**

$$M_X(t)=\sum_{x=0}^{\infty} e^{tx}\,P(X=x)=\sum_{x=0}^{\infty} e^{tx}\left\{a_x\,\theta^x/f(\theta)\right\}$$

$$=\frac{1}{f(\theta)}\sum_{x=0}^{\infty} a_x\,(\theta e^t)^x=\frac{f(\theta e^t)}{f(\theta)} \qquad \ldots(7\cdot35)$$

**7·9·2. Recurrence Relation for Cumulants of p.s.d.** The cumulant generating function is given by

$$K_X(t)=\log M_X(t)=\log\left[\frac{f(\theta.e^t)}{f(\theta)}\right]$$

$$\sum_{r=1}^{\infty}\kappa_r\,\frac{t^r}{r!}=\log f(\theta\,e^t)-\log f(\theta) \qquad \ldots(1)$$

Differentiating (1) partially w.r.to $\theta$ and $t$ respectively, we get

$$\sum_{r=1}^{\infty} \frac{\partial}{\partial \theta} \kappa_r \frac{t^r}{r!} = \frac{e^t f'(\theta e^t)}{f(\theta e^t)} - \frac{f'(\theta)}{f(\theta)} \quad \text{...(2)}$$

and
$$\sum_{r=1}^{\infty} \kappa_r \frac{r\, t^{r-1}}{r!} = \frac{\theta e^t f'(\theta e^t)}{f(\theta e^t)} \quad \text{....(3)}$$

Subtracting (3) from $\theta$ times (2), we get

$$\theta \sum_{r=1}^{\infty} \frac{\partial}{\partial \theta} \kappa_r \frac{t^r}{r!} = \sum_{r=1}^{\infty} \kappa_r \cdot \frac{t^{r-1}}{(r-1)!} - \frac{\theta \cdot f'(\theta)}{f(\theta)}$$

Comparing like powers of $t$ on both sides, we get

$$0 = \kappa_1 - \frac{\theta f'(\theta)}{f(\theta)} \Rightarrow \kappa_1 = \frac{\theta f'(\theta)}{f(\theta)} \quad \text{...(7·36)}$$

and $\kappa_{r+1} = \theta \cdot \dfrac{d}{d\theta} \kappa_r ;\ r = 1, 2, \ldots$ (Comparing co-efficient of $t^r/r!$)

**Remark.** We have

$$\text{Mean} = \kappa_1 = \frac{\theta f'(\theta)}{f(\theta)}, \quad \text{...(7·36 } a\text{)}$$

Alternatively

$$\text{Mean} = \sum_{x=0}^{\infty} x \left\{ a_x \theta^x / f(\theta) \right\} = \frac{\theta}{f(\theta)} \sum_{x=0}^{\infty} x\, a_x\, \theta^{x-1} = \frac{\theta f'(\theta)}{f(\theta)}$$

**7·9·3. Particular Cases of g.p.s.d. 1.** *Binomial Distribution.* Let us take $\theta = p/(1-p)$, $f(\theta) = (1+\theta)^n$ and $S = \{0, 1, 2, \ldots, n\}$, a set of $(n+1)$ non-negative integers then

$$f(\theta) = \sum_{x \in S} a_x \theta^x \Rightarrow (1+\theta)^n = \sum_{x=0}^{n} a_x \theta^x$$

$$\Rightarrow \qquad a_x = \binom{n}{x}$$

$$\therefore P(X = x) = \frac{\binom{n}{x} \left[ \dfrac{p}{(1-p)} \right]^x}{\left[ 1 + \dfrac{p}{(1-p)} \right]^n}$$

$$= \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} ; & x = 0, 1, \ldots, n \\ 0, & \text{otherwise} \end{cases}$$

which is the probability mass function of the binomial distribution with parameters $n$ and $p$.

**2. Negative Binomial Distribution.** Let us take $\theta = p/(1+p)$, $f(\theta) = (1-\theta)^{-n}$ and $S = \{0, 1, 2, \ldots \text{ ad infinity}\}$, $0 \le \theta < 1$, $n > 0$.

Now $f(\theta) = \sum_{x \in S} a_x \theta^x \Rightarrow (1-\theta)^{-n} = \sum_{x=0}^{\infty} a_x \theta^x$

$$\Rightarrow \qquad a_x = (-1)^x \binom{-n}{x} = (-1)^x \cdot (-1)^x \binom{n+x-1}{x} = \binom{n+x-1}{x}$$

$$\therefore \quad P(X=x) = \sum_{x=0}^{\infty} \binom{n+x-1}{x} \left[(p/1+p)\right]^x \Big/ \left[1 - \{p/(1+p)\}\right]^{-n}$$

$$= \sum_{x=0}^{\infty} \binom{n+x-1}{x} p^x (1+p)^{-(n+x)};\ x = 0, 1, 2, \ldots$$

$$= \sum_{x=0}^{\infty} \binom{-n}{x} (1+p)^{-(n+x)} (-p)^x;\ x = 0, 1, 2, \ldots$$

which is the probability mass function of the negative binomial distribution.

**3. Logarithmic Series Distribution.** Let $f(\theta) = -\log(1-\theta)$ and $S = \{1, 2, 3, \ldots\}$.

Then $f(\theta) = \sum_{x \in S} a_x \theta^x \Rightarrow -\log(1-\theta) = \sum_{x=1}^{\infty} a_x \theta^x$, *i.e.*, $a_x = \dfrac{1}{x}$

$$\therefore P(X=x) = \begin{cases} \dfrac{a_x \theta^x}{f(\theta)} = \dfrac{\theta^x}{x[-\log(1-\theta)]}; & = 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}$$

**4. Poisson Distribution.** Let $f(\theta) = e^{\theta}$, $S = \{0, 1, 2, \ldots\}$. Then

$$f(\theta) = \sum_{x \in S} a_x \theta^x \Rightarrow e^{\theta} = \sum_{x=0}^{\infty} a_x \theta^x \quad i.e., \quad a_x = \frac{1}{x!}$$

$$\therefore \quad P(X=x) = \frac{a_x \theta^x}{f(\theta)} = \frac{\theta^x}{x!\,e^{\theta}} = \frac{e^{-\theta}\theta^x}{x!};\ x = 0, 1, 2, \ldots$$

which is the probability mass function of the Poisson distribution with parameter $\theta$.

## ADDITIONAL EXERCISES ON CHAPTER VII

1. Show that the necessary and sufficient conditions for two given numbers $a, b$ to be respectively the mean and the variance of some binomial distribution are that $a > b > 0$ and $\dfrac{a^2}{a-b}$ is an integer.

Show further that when these conditions are satisfied, the binomial distribution is uniquely determined.

2. In a game of taking a chance, a contestant has to give correct answers to 4 out of 5 questions to win the contest. Questions are given with 3 answers each, out of which one is a correct answer. If a contestant answers the questions by selecting the answers at random, what is the probability that he will win the contest?

**Ans.** $10/3^5 = 0 \cdot 0412$

3. Suppose the automatic machines of a plant fail with probability $q$, the machine failure is independent from machine to machine and the plant stays in operation, if at least half of the machines run. Consider a two-machine plant and a four-machine plant. Show that the value of $q$ for uninterrupted operations,

(*i*) when the value of $q$ is same in both plants is $\frac{1}{2}$,

(*ii*) when a two-machine plant is perferred is $q > \frac{1}{2}$, and

(*iii*) when a four-machine plant is preferred is $q < \frac{1}{2}$

4. If $b(r; n, p) = {}^nC_r\, p^r \cdot q^{n-r}$ is the binomial probability in the usual notation and if $B(k; n, p) = \sum_{r=0}^{k} b(r; n, p)$, prove the following results for the "tails" of the binomial distribution.

(*i*) $1 - B(k-1; n, p) \leq \dfrac{n}{k - np}\, b(k; n, p), \; k > np + 1$

(*ii*) $B(k; n, p) \leq \dfrac{n}{np - k}\, b(k; n, p), k < np$

(*iii*) $1 - B(k; n, p) = n \dbinom{n-1}{k} \int_0^p t^k (1-t)^{n-k-1}\, dt$

5. If a coin is tossed $n$ times where $n$ is very large even number, show that the probability of getting exactly $(\frac{1}{2} n - p)$ heads and $(\frac{1}{2} n + p)$ tails is approximately

$$\left(\frac{2}{\pi n}\right)^{\frac{1}{2}} e^{-2p^2/}$$

6. If $X \sim B(n, p)$, show that $P(X \leq 2) = P[X \geq (n-2)]$, if and only if $p = \frac{1}{2}$.

[Calcutta Univ. B.Sc. (Hons.), 1989]

7. If $X \sim B(n, p)$, show that $X$ is symmetrically distributed about $c$ if and only if $p = 1/2$ and $c = n/2$. (Madurai Univ. M.A., 1991)

8. If $X \sim B(n, p)$, and $Y = X^2$, find corr. $(X, Y)$

[Delhi Univ. (Stat Hons.) Spl. Course; 1989]

9. $A$ and $B$ have equal chances of winning a single game, $A$ wants $n$ games and $B$, $n + 1$ games to win a match. Show that the odds in favour of $A$ are $1 + P$ to $1 - P$, where $P = \dfrac{(2n)!}{n!\, n!\, 2^{2n}}$

**Hint.** The probability that $A$ wins at least $n$ games is

$${}^{2n}C_n q^n p^n + {}^{2n}C_{n+1} q^{n-1} p^{n+1} + \ldots + {}^{2n}C_{2n} p^{2n}$$

Now ${}^{2n}C_0 + {}^{2n}C_1 + \ldots + {}^{2n}C_{n-1} + {}^{2n}C_n + {}^{2n}C_{n+1} + \ldots + {}^{2n}C_{2n} = 2^{2n}$

$\therefore\ {}^{2n}C_{n+1} + {}^{2n}C_{n+3} + \ldots + {}^{2n}C_{2n} = \frac{1}{2}[2^{2n} - {}^{2n}C_n]$

$\therefore$ Probability of A's win $= \dfrac{1}{2^{2n}}\left(\frac{1}{2}(2^{2n} - {}^{2n}C_n)\right) = \frac{1}{2}(1 - P)$

$\therefore$ Probability of A's losing $= 1 - \frac{1}{2}(1 - P) = \frac{1}{2}(1 + P)$

Hence the result.

10. (*a*) The chance of success in each Bernoulli trial is $p$. If $p_k$ is the probability that there are even number of successes in $k$ trials, prove that

$$p_k = p + p_{k-1}(1 - 2p) \qquad \ldots(*)$$

Deduce that $p_k = \dfrac{1}{2}\left[1 + (1 - 2p)^k\right]$

(*b*) Also obtain the probability generating function of (*) and hence obtain an explicit expression for $p_k$

(*c*) Obtain an expression for $p_k$ directly without using (*a*) or (*b*).

11. A spider and a fly are situated at the corners (0, 0) and $(n, n)$ of a rectangular grid. The spider walks only north or east, the fly only south or west. They take their steps simultaneously to an adjacent vertex of the grid. Show that, if the successive steps are independent and equally likely to go in each of the two possible directions, the probability that they will meet is $\dbinom{2n}{n}\left(\dfrac{1}{2}\right)^{2n}$

[Delhi Univ. B.Sc. (Statistics Hons.) Spl. Course, 1988]

12. For the binomial distribution, show that the probability that the number of successes in $n$ trials should not exceed $x$ is given by

$$\frac{\displaystyle\int_{p/q}^{\infty} \frac{y^x}{(1 + y)^{n+1}}\, dy}{\displaystyle\int_{0}^{\infty} \frac{y^x}{(1 + y)^{n+1}}\, dy}, \quad p + q = 1$$

where $p$ is the probability of success.

13. Prove the identity

$$p^n + \binom{n}{1} p^{n-1} q + \binom{n}{2} p^{n-2} q^2 + \ldots + \binom{n}{k} p^{n-k} q^k = \sum_{x=0}^{k} \binom{n}{x} \left(p^{n-x} q^x\right)$$

$$= \frac{\displaystyle\int_0^p x^{n-k-1} (1 - x)^k\, dx}{\displaystyle\int_0^1 x^{n-k-1} (1 - x)^k\, dx}$$

**Hint.** For Questions 12 and 13, see Example 7·23.

**14.** Let $X$ be a random variable whose probability function is $b(x; n, p)$. Let $Y = X/n$ be a new random variable. Show that the expected value of $Y$ is $p$ and the variance of $Y$ is $pq/n$. If $p(y)$ is the probability function for $Y$, show that $p(y) = b(ny; n, p)$. What are the possible values that $y$ can take on ?

**15.** Suppose that the number of telephone calls that an operator receives from 9·00 to 9·05 hours in a day follows a Poisson distribution with mean 3. Find the probability that (*i*) the operator will receive no calls in that time interval tomorrow, (*ii*) In the next three days the operator will receive a total of 1 call in that time interval.

**Ans.** (*i*) $e^{-3}$ (*ii*) $3 \times (e^{-3})^2 (1 - e^{-3})$.

**16.** A large number of observations on a given solution which contained bacteria were made taking samples 1 ml. each, noting down the number of bacteria present in each sample. Assuming the Poisson distribution, and given that 10% samples contained no bacteria, find the average number of bacteria per ml.

**Ans.** $\log_e 10$ or 2·3026

**17.** The number of oil tankers, say $N$, arriving at a certain refinery each day has a Poisson distribution with parameter 2. Present port facilities can service three tankers a day. If more than three tankers arrive in a day, the tankers in excess of three must be sent to another port.

(*i*) On a given day, what is the probability of having to send tankers away?

(*ii*) How much must present facilities be increased to permit handling all tankers on approximately 90 per cent of days ?

(*iii*) What is the expected number of tankers arriving per day ?

(*iv*) What is the expected number of tankers serviced daily? and

(*v*) What is the expected number of tankers turned away daily?

**Ans.** (*i*) 0·145, (*ii*) 4, (*iii*) 2, (*iv*) 1·785 and (*v*) 0·215.

**18.** If $X$ is any non-negative integer valued variate and $a$ is any positive number, show that

$$P(X \geq a) \leq t^{-a} . E(t^X) ; t > 1$$

Verify the inequality

$$P(X \geq 2\lambda) \leq (e/4)^{\lambda} \text{ when } X \sim P(\lambda).$$

**19.** If $X$ is any non-negative integer valued and $a$ is any positive number, show that

$$P(X \leq a) \leq t^{-a} E(t^X), \; 0 < t \leq 1.$$

Verify the inequality :

$$P(X \leq \frac{1}{2} m) \leq (2/e)^{m/2}, \text{ when } X \sim P(m)$$

**20.** Suppose $(X, Y)$ have the joint p.m.f.

$$f(x, y) = \frac{e^{-(a+b)} a^x b^{y-x}}{x!(y-x)!}, x = 0, 1, 2, \ldots ; y = x, x+1, \ldots$$

Show that the correlation coefficient between $X$ and $Y$ is $[a/(a+b)]^{1/2}$.

Also obtain the distribution of $Y - X$.

[Delhi Univ. (Spl. Course. Statistics Hons.), 1988]

**Hint.** $M_{X,Y}(t_1, t_2) = \sum\limits_{x=0}^{\infty} \sum\limits_{y=x}^{\infty} \left[ e^{t_1 x + t_2 y} \dfrac{e^{-(a+b)} \cdot a^x b^{y-x}}{x!(y-x)!} \right]$

$$= e^{-(a+b)} \left[ \sum_{x=0}^{\infty} \frac{\left(a e^{t_1} e^{t_2}\right)^x}{x!} \sum_{z=0}^{\infty} \frac{\left(b e^{t_2}\right)^z}{z!} \right]; (y - x = z)$$

$$= \exp\left[a e^{t_1 + t_2} + b e^{t_2} - a - b\right] \qquad \ldots(**)$$

$$M_X(t_1) = M(t_1, 0) = \exp\left[a(e^{t_1} - 1)\right] \Rightarrow X \sim P(a)$$

$$M_Y(t_1) = M(0, t_2) = \exp\left[(a+b)\left\{e^{t_2} - 1\right\}\right] \Rightarrow Y \sim P(a+b)$$

$$E(XY) = \frac{\partial^2 M(t_1, t_2)}{\partial t_1 \, \partial t_2}\bigg|_{t_1 = t_2 = 0} = a^2 + ab + a$$

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = a^2 + ab + a - a(a+b) = a$$

**Distribution of Y − X.** Taking $t_1 + t_2 = 0 \Rightarrow t_1 = -t_2$ in (**),

$$M(-t_2, t_2) = E(e^{-t_2 X + t_2 Y}) = E(e^{(Y-X)t_2}) = \exp\left[b(e^{t_2} - 1)\right]$$

$$\Rightarrow \qquad Y - X \sim P(b)$$

**21.** If $X$ is Negative Binomial variate with parameters ($k$ and $Q^{-1}$), prove that

$$P(X \ge m) = \frac{1}{B(m,k)} \cdot \int_0^P \frac{x^{m-1} dx}{(1+x)^{k+m}}; \; Q - P = 1$$

**Hint.** $P(X \ge m) = \sum\limits_{r=m}^{\infty} \dbinom{-k}{r} Q^{-k-r}(-P)^r$

$$= \sum_{r=m}^{\infty} \binom{k+r-1}{r} Q^{-k-r} P^r$$

$$\frac{d}{dP}[P(X \ge m)] = \sum_{r=m}^{\infty} (T_r - T_{r+1}); \; T_r = r \cdot \binom{k+r-1}{r} P^{r-1} Q^{-k-r}$$

$$= T_m = \frac{1}{B(m,k)} \cdot P^{m-1} Q^{-k-m}$$

Integrating, we get

$$P(X \geq m) = \frac{1}{B(m,k)} \int_0^P \frac{x^{m-1}\, dx}{(1+x)^{k+m}} \qquad (\because Q - P = 1)$$

(b) If $X$ is N.B. $(k, p)$, show that

$$P(X \geq m) = \frac{1}{B(m,k)} \cdot \int_p^1 y^{k-1} (1-y)^{m-1}\, dy$$

**22.** In a sequence of independent trials, the probability of a success on each trial is '$p$'. By considering the outcome of the first trial, show that $G_r(t)$, the p.g.f. of the number of trials required to achieve the $r$th success, satisfies :

$$G_r(t) = pt\, G_{r-1}(t) + qt\, G_r(t)$$

and hence obtain $G_r(t)$. [Delhi Univ. (Spl. Course Statistics Hons.), 1987]

Ans. $G_r(t) = [pt/(1-qt)]^r$

**23.** Let $X$ and $Y$ be independent random variables with the same (geometric) distribution given by $P(X = k) = pq^k$; $k = 0, 1, 2, ...$

Let $Z = \max(X, Y)$

*(i)* Find the probability distribution of $Z$.

*(ii)* Find the joint probability distribution of $X$ and $Z$.

*(iii)* Find the conditional probability distribution of $X$ given $Z = l$, *i.e.*, compute $P(X = k \mid Z = l)$ for *all* $k, l = 0, 1, 2, ...$

*(iv)* Find the conditional probability distribution of $Z$ given $X = k$, i.e. compute $P(Z = l \mid X = k)$ for all $k, l = 0, 1, 2, ...$

Ans. (i) $P(Z = l) = pq^l\left[2 - q^l - q^{l+1}\right]$; $l = 0, 1, 2, ...$

(ii) $$P(Z = l \cap X = k) = \begin{cases} 0, \text{ if } l < k \\ p\, q^k (1 - q^{k+1}) \text{ if } l = k = 0, 1, 2, ... \\ p^2 q^{k+1} \text{ if } l > k = 0, 1, 2, ... \end{cases}$$

(iii) $$P(X = k \mid Z = l) = \begin{cases} 0 \text{ if } l < k \\ (1 - q^{k+1})/(2 - q^k - q^{k+1}) \text{ if } l = k = 0, 1, 2, ... \\ pq^k/(2 - q^l - q^{l+1}) \text{ if } l > k = 0, 1, 2, ... \end{cases}$$

(iv) $$P(Z = l \mid X = k) = \begin{cases} 0 \text{ if } l < k \\ 1 - q^{k+1} \text{ if } l = k = 0, 1, 2, \dots \\ pq^l \text{ if } l > k = 0, 1, 2, \dots \end{cases}$$

**24.** Suppose that $X_1, X_2, ..., X_n$ are mutually independent indicator random variables, with $P(X_i = 1) = p$, $0 < p < 1$. Show that for $1 \leq M \leq N$,

$$P\left(\sum_{i=1}^{M} X_i = k \,\middle|\, \sum_{i=1}^{N} X_i = n\right) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

**25.** Suppose one makes $(m + n)$ independent trials of an experiment whose probability of success at each trial is $p$. Let $q = (1 - p)$. Show that for any $k = 0, 1, 2, \dots n$, the conditional probability that exactly $(m + k)$ trials will result in success, given that the first $m$ trials result in success, is equal to $\binom{n}{k} p^k q^{n-k}$. Show further that the conditional probability that exactly $(m + k)$ trials will result in success. given that at least $m$ trials result in success is equal to

$$\binom{m+n}{m+k}\left(\frac{p}{q}\right)^k \Big/ \sum_{r=0}^{n} \binom{m+n}{m+r}\left(\frac{p}{q}\right)^r$$

**26.** Let $X_1, X_2, \dots, X_n$ be independent Bernoulli variates with common parameter $p = P(X_1 = 1)$. Let $S_j = X_1 + X_2 + \dots + X_j$ for $1 \le j \le n$. Show that $P(S_j = r \mid S_n = s)$ does not depend on $p$ $(0 < p < 1)$ and takes the form of a hypergeometric probability for $1 \le j \le n$, $0 \le r \le s \le n$.

**Hint.** $S_n \sim B(n, p)$

$$P(S_j = r \cap S_n = s) = \binom{j}{r} p^r q^{j-r} \cdot \binom{n-j}{s-r} p^{s-r} q^{n-j-s+r}$$

$$P(S_j = r \mid S_n = s) = P(S_j = r \cap S_n = s)/P(S_n = s)$$

$$= \binom{j}{r}\binom{n-j}{s-r} \div \binom{n}{s}; \; 1 \le j \le n, 0 \le r \le s \le n$$

a result, which is independent of $p$.

**27.** An urn contains $w$ white balls and $b$ black balls. Balls are drawn one at a time from the urn, without replacement. Find the distribution of the number $X$ of draws needed to obtain the $k$ th black ball. Find also the factorial moments $E\left[X^{(r)}\right]$. **[Delhi Univ. B.Sc. Statistics Hons. (Spl. Course), 1989]**

**Ans.** $$P(X = x) = \frac{\binom{b}{k-1}\binom{w}{x-k}}{\binom{b+w}{x-1}} \times \left(\frac{b - k + 1}{b + w - x + 1}\right)$$

$$= \binom{x-1}{k-1}\binom{b+w-x}{b-k} \div \binom{b+w}{b}; \text{ ( On simplification )}$$

For $E\left[X^{(r)}\right]$, proceed as in § 7·6·2.

$$E\left[X^{(r)}\right] = k^{(r)} (b + w + 1)^{(r)}/(b + 1)^{(r)}$$

**28.** The joint p.m.f. of two discrete r.v.'s $X_1$ and $X_2$ is:

$$p(x_1, x_2) = \binom{n_1}{x_1}\binom{n_2}{x_2 - x_1} p^{x_2} (1 - p)^{n_1 + n_2 - x_2}$$

with $x_1 \le x_2 \le n_2 + x_1$; $0 \le x_1 \le n_1$.

Find the marginal distributions of $X_1$ and $X_2$.

**Ans.** $X_1 \sim B\,(n_1, p)$ and $X_2 \sim B\,(n_1 + n_2, p)$

**29.** Two discrete random variables $X$ and $Y$ have the joint probability distribution :

$$p(x, y) = \frac{9\,!}{x\,!\,y\,!\,(9 - x - y)\,!}\left(\frac{1}{3}\right)^9, \text{ where}$$

$$0 \le x \le 9,\ 0 \le y \le 9 \text{ and } 0 \le (x + y) \le 9$$

(*i*) Show that the marginal distribution of $X$ is binomial with parameters 9 and ⅓ .

(*ii*) Show that the conditional distribution of $Y$ given $X = 3$ is also binomial with parameters 6 and ½ .

**30.** A Polya process is defined by the quantities :

$$P_k(t) = \left[\frac{\lambda t}{1 + b\lambda t}\right]^k \frac{1(1 + b)\ldots\{1 + (k - 1)b)\}}{k\,!}\,P_0(t)$$

where $P_0(t) = (1 + b\lambda t)^{-1/b}$

and $\lambda$, $b$ are parameters, $t$ is a continuous variable and $k$ may take zero or poisitive integral values only. Verify that the distribution satisfies the requirements for a probability distribution in $K$ and find the expectation of $K$ and its variance.

**Hint.** Let $K$ be a random variable with the distribution,

$P(K = k) = P_k(t)\,;\ k = 0, 1, 2, \ldots, \infty.$

$$\sum_{k=0}^{\infty} P_k(t) = (1 + b\lambda t)^{-1/b} \sum_{k=0}^{\infty}\left[\frac{\lambda t}{1 + b\lambda t}\right]^k \times \frac{b^k\left[\frac{1}{b}\right]\left[\frac{1}{b} + 1\right]\ldots\left[\frac{1}{b} + k - 1\right]}{k\,!}$$

$$= (1 + b\lambda t)^{-1/b} \sum_{k=0}^{\infty}\left[\frac{\lambda t}{1 + b\lambda t}\right]^k b^k \binom{(1/b) + k - 1}{k}$$

$$= (1 + b\lambda t)^{-1/b}\,\frac{1}{\left[1 - \dfrac{b\lambda t}{1 + b\lambda t}\right]^{1/b}}\ * = 1$$

$\therefore P_k(t)$ represents a probability distribution for every fixed $b$, $\lambda$ and $t$.

M.G.F. of $K = E(e^{Ku}) = \sum\limits_{k=0}^{\infty} e^{uk} P_k(t)$.

---

$$*\,(1 - x)^{-n} = \sum_{k=0}^{\infty}\binom{-n}{k}(-x)^k = \sum_{k=0}^{\infty}\binom{n + k - 1}{k}x^k$$

$$= (1+b\lambda t)^{-1/b} \sum_{k=0}^{\infty} \left[ e^{ku} \frac{(\lambda t)^k}{(1+b\lambda t)^k} \frac{b^k \left[\frac{1}{b}\right]\left[\frac{1}{b}+1\right]\cdots\left[\frac{1}{b}+k-1\right]}{k!} \right]$$

$$= (1+b\lambda t)^{-1/b} \sum_{k=0}^{\infty} \left[\frac{e^u b\lambda t}{1+b\lambda t}\right]^k \binom{(1/b)+k-1}{k}$$

$$= (1+b\lambda t)^{-1/b} \frac{1}{\left[1-\frac{e^u b\lambda t}{1+b\lambda t}\right]^{1/b}} = [1+b\lambda t - e^u b\lambda t]^{-1/b}$$

$$= g(u), \text{ (say)}$$

$$\therefore \quad g'(u) = \frac{1}{b}\left[1+b\lambda t - e^u b\lambda t\right]^{-(1/b)-1} \cdot (e^u b\lambda t)$$

$$E(K) = \mu_1' \text{ (about origin)} = \text{Mean} = [g'(u)]_{u=0}$$

$$= \left[\frac{1}{b}\right]\left[1+b\lambda t - b\lambda t\right]^{-(1/b)-1} \cdot (b\lambda t) = \lambda t$$

Similarly $\quad \mu_2' = [g''(u)]_{u=0} = (b+1)\lambda^2 t^2 + \lambda t$

$\therefore \quad$ Variance $= \mu'_2 - \mu_1'^2 = \lambda t\,(1+b\lambda t)$.

## OBJECTIVE TYPE QUESTIONS

1. (*i*) Match the correct parts to make a valid statement :

| | | | |
|---|---|---|---|
| (*a*) | Binomial distribution applies to | 1. | rare events |
| (*b*) | Poisson distribution applies to | 2. | repeated two alternatives. |
| (c) | The mean of a Hypergeometric distribution | 3. | $\frac{1-6pq}{npq}$ |
| (*d*) | The moment generating function of negative binomial distribution | 4. | $n \cdot \frac{M}{n}\left(1-\frac{M}{n}\right)\left(\frac{N-n}{N-1}\right)$ |
| (*e*) | The coefficient of kurtosis of a binomial distribution | 5. | $(Q - pe^t)^{-r}$ |
| (*f*) | The variance of geometric distribution | 6. | $\frac{nM}{N}$ |
| (*g*) | Variance of Hypergeometric distribution | 7. | $\frac{q}{p^2}$ |

**II.** Under what conditions binomial distribution tends to (*i*) Poisson distribution, (*ii*) Normal distribution, (*iii*) Geometric distribution. Give practical examples (one each) where you would expect binomial, Poisson, negative binomial and geometric distribution.

**III.** State the relationship between :

(*i*) Mean and variance of Poisson distribution.

(*ii*) Mean and variance of negative binomial distribution.

(*iii*) Mean and variance of geometric distribution.

(*iv*) Poisson distribution and binomial distribution.

(*v*) Hypergeometric distribution and binomial distribution.

**IV.** Name the discrete distribution for which

(*i*) Mean and variance have the same value.

(*ii*) Mean is greater than the variance.

**V.** State which of the following statements are *True* and which are *False*. In case of the false statement, give the correct statement :

(*i*) Mean of binomial distribution is 3 and variance is 5.

(*ii*) Mean of Poisson distribution is 2 and variance is 3.

(*iii*) The sum of two independent Poisson variates is also a Poisson variate. The result holds for the difference also.

(*iv*) For a binomial distribution,

Mean = Mode = Median

(*v*) The Poisson distribution is a limiting case of binomial distribution when $n \to \infty, p \to 0, np \to m$.

(*vi*) Nearly all the distributions are particular cases of Poisson distribution.

(*vii*) The sum of two binomial variates is a binomial variate if the variables are independent and have the different probabilities of success.

(*viii*) Negative binomial distribution may be regarded as the generalisation of geometric distribution.

**VI.** Fill in the blanks :

(*i*) The variance of a binomial distribution is .........

(*ii*) The $\beta$-coefficient of skewness of the binomial distribution is .........

(*iii*) The moment generating function of Poisson distribution is .........

(*iv*) The characteristic function of negative binomial distribution is .........

(*v*) The coefficient of skewness of a Poisson distribution is .........

(*vi*) Poisson distribution is a limiting case of binomial distribution under the conditions .........

(*vii*) For Poisson distribution all cumulants ...........

(*viii*) Mean > variance for ......... distribution.

(*ix*) For the Poisson distribution, the variance and the third central moment are ...........

(*x*) Mean < variance for ......... distribution.

**VII.** Give the correct answer to each of the following :

(*i*) The skewness in a binomial distribution will be zero, if
(*a*) $p < \frac{1}{2}$, (*b*) $p = \frac{1}{2}$, (*c*) $p > \frac{1}{2}$, (*d*) $p < q$.

(*ii*) The mean and variance of negative binomial distribution :
(*a*) are same, (*b*) cannot be same, (*c*) are sometimes equal in limiting case, as $n \to \infty$.

(*iii*) The characteristic function of Poisson distribution $P(m)$ is
(*a*) $e^{m(it-1)}$, (*b*) $e^{m(e^{it}-1)}$, (*c*) $e^{mit}$, (*d*) none of these.

(*iv*) The coefficient of variation of Poisson distribution with mean 4 is
(*a*) $\frac{1}{4}$, (*b*) $\frac{2}{4}$, (*c*) 4, (*d*) 2

(*v*) The coefficient of kurtosis of a Poisson distribution with mean $m$ is
(*a*) $1/m$, (*b*) $-1/m$, (*c*) $m$, (*d*) $3 + (1/m)$

(*vi*) The mean of a Hypergeometric distribution is
(*a*) $\dfrac{N(M-1)}{N(N-1)}$, (*b*) $\dfrac{M(M-1)}{N(N-1)}$, (*c*) $\dfrac{NM(M-1)}{N(N-1)}$, (*d*) None of these

(*vii*) In a Poisson distribution, the second moment about the origin is 12. Then its third moment about mean is (*a*) 2, (*b*) 3, (*c*) 5, (*d*) 10.

(*viii*) The mean of the binomial distribution ${}^{10}C_x \left(\frac{2}{5}\right)^x \left(\frac{3}{5}\right)^{10-x}$ ; $x = 0, 1, 2, \ldots, 10$ is (*a*) 4, (*b*) 6, (*c*) 5, (*d*) 0.

(*ix*) The mean of Poisson variate is
(*a*) greater than, (*b*) less than, (*c*) equal to, (*d*) twice, its variance.

(*x*) The moment generating function of Geometric distribution is
(*a*) $p(1-qe^t)$, (*b*) $p/(1-qe^t)$, (*c*) $pe^t/(1-qe^t)$, (*d*) None of these.

**VIII.** By using the uniqueness property of m.g.f.'s, determine the distribution if the M.G.F. is as follows :

(*a*) $M(t) = \left(\frac{1}{2} + \frac{1}{2}e^t\right)^6$, (*b*) $M(t) = \dfrac{(1+e^t)^5}{32}$

(*c*) $M(t) = \dfrac{(1+2e^t)^3}{27}$, (*d*) $M(t) = e^{3(e^t-1)}$,

(*e*) $M(t) = e^{(e^t-1)/4}$, (*f*) $M(t) = \frac{1}{3}e^{-t}\left(e^{-t} - \frac{2}{3}\right)^{-1}$

(*g*) $M(t) = 4(3e^{-t}-1)^{-2}$, (*h*) $M(t) = (3e^{-t}-2)^{-3}$

**Ans.** (*a*) Binomial, $n = 6, p = \frac{1}{2}$ ; (*b*) Binomial, $n = 5, p = \frac{1}{2}$ ;
(*c*) Binomial, $n = 3, p = \frac{2}{3}$ ; (*d*) Poisson, $\lambda = 3$,
(*e*) Poisson, $\lambda = \frac{1}{4}$, (*f*) Geometric with $p = 1/3$,
(*g*) Negative binomial with $r = 2, p = \frac{2}{3}$ ;
(*h*) Negative binomial with $r = 3, p = 1/3$.

CHAPTER EIGHT

# *Theoretical Continuous Distributions*

**8·1. Rectangular (or Uniform Distribution.** A random variable $X$ is said to have a continuous uniform distribution over an interval $(a, b)$ if its probability density function is constant $= k$ (say), over the entire range of $X$, *i.e.*,

$$f(x) = \begin{cases} k, a < x < b \\ 0, \text{ otnerwise} \end{cases}$$

Since total probability is always unity, we have

$$\int_a^b f(x)\,dx = 1 \quad \Rightarrow \quad k\int_a^b dx = 1 \quad i.e., \quad k = \frac{1}{b-a}$$

$$\therefore \qquad f(x) = \begin{cases} \dfrac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases} \qquad \text{...(8·1)}$$

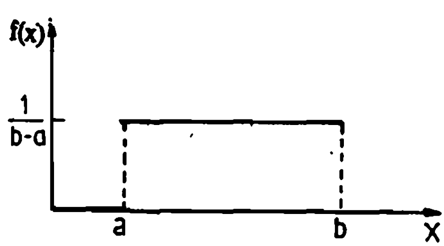
**Remarks. 1.** $a$ and $b$, $(a < b)$ are the two parameters of the uniform distribution on $(a, b)$.

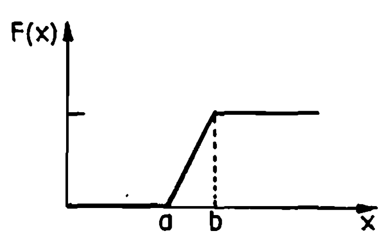
**2.** The distribution is also known as rectangular distribution, since the curve $y = f(x)$ describes a rectangle over the $x$-aixs and between the ordinates at $x = a$ and $x = b$.

**3.** The distribution function $F(x)$ is given by

$$F(x) = \begin{cases} 0, \text{ if } -\infty < x < a \\ \dfrac{x-a}{b-a}, \; a \le x \le b \\ 1, \quad b < x < \infty \end{cases} \qquad \text{...(8·1 } a\text{)}$$

Since $F(x)$ is not continuous at $x = a$ and $x = b$, it is not differentiable at these points. Thus $\dfrac{d}{dx} F(x) = f(x) = \dfrac{1}{b-a} \neq 0$, exists everywhere except at the points $x = a$ and $x = b$. and consequently p.d.f. $f(x)$ is given by (8·1).



4. The graphs of uniform p.d.f. $f(x)$ and the corresponding distribution function $F(x)$ are given on page 8·1 :

5. For a rectangular or uniform variate $X$ in $(-a, a)$, p.d.f. is given by

$$f(x) = \begin{cases} \dfrac{1}{2a}, & -a < x < a \\ 0, & \text{otherwise.} \end{cases}$$

**8·1·1. Moments of Rectangular Distribution.**

$$\mu'_r = \int_a^b x^r f(x)\, dx = \frac{1}{(b-a)} \int_a^b x^r\, dx = \frac{1}{(b-a)} \left[ \frac{b^{r+1} - a^{r+1}}{r+1} \right] \qquad \ldots(8\cdot2)$$

In particular

$$\text{Mean} = \mu_1' = \frac{1}{(b-a)} \left[ \frac{b^2 - a^2}{2} \right] = \frac{b+a}{2}$$

and $$\mu_2' = \frac{1}{(b-a)} \left[ \frac{b^3 - a^3}{3} \right] = \frac{1}{3}(b^2 + ab + a^2)$$

$$\therefore \qquad \mu_2 = \mu_2' - \mu_1'^2 = \frac{1}{3}(b^2 + ab + a^2) - \left( \frac{b+a}{2} \right)^2 = \frac{1}{12}(b-a)^2$$

**8·1·2. Moment Generating Function** is given by

$$M_X(t) = \int_a^b e^{tx} f(x)\, dx = \frac{e^{bt} - e^{at}}{t(b-a)}$$

**8·1·3. Characteristic Function** is given by

$$\varphi_X(t) = \int_a^b e^{itx} f(x)\, dx = \frac{e^{ibt} - e^{iat}}{it(b-a)}$$

**8·1·4. Mean Deviation about Mean,** $\eta$ is given by

$$\eta = E\,|X - \text{Mean}| = \int_a^b |x - \text{Mean}|\, f(x)\, dx$$

$$= \frac{1}{(b-a)} \int_a^b \left| x - \frac{a+b}{2} \right| dx$$

$$= \frac{1}{(b-a)} \int_{-(b-a)/2}^{(b-a)/2} |t|\, dt \qquad \left[ t = x - \frac{a+b}{2} \right]$$

$$= \frac{1}{(b-a)} \cdot 2 \int_0^{(b-a)/2} t\, dt = \frac{b-a}{4}$$

**Example 8·1** *If $X$ is uniformly distributed with mean 1 and variance 4/3, find $P(X < 0)$.* **[Delhi Univ. B.A. (Hons. Spl. Course-Statistics), 1989]**

**Solution.** Let $X \sim U[a, b]$, so that $p(x) = \dfrac{1}{b-a}$ ; $a < x < b$. We are given:

$$\text{Mean} = \frac{b+a}{2} = 1 \quad \Rightarrow \quad b + a = 2$$

$$\text{Var}(X) = \frac{1}{12}(b-a)^2 = \frac{4}{3} \quad \Rightarrow \quad (b-a)^2 = 16 \quad \Rightarrow \quad b - a = \pm 4$$

Solving, we get : $a = -1$ and $b = 3$ ; $(a < b)$.

$$\therefore \qquad p(x) = \frac{1}{4} \; ; \; -1 < x < 3$$

$$P(X < 0) = \int_{-1}^{0} p(x)\, dx = \frac{1}{4} \mid x \mid_{-1}^{0} = \frac{1}{4}$$

**Example 8·2.** *Subway trains on a certain line run every half hour between mid-night and six in the morning. What is the probability that a man entering the station at a random time during this period will have to wait at least twenty minutes?*

**Solution.** Let the *r.v.* $X$ denote the waiting time (in minutes) for the next train. Under the assumption that a man arrives at the station at random, $X$ is distributed uniformly on (0, 30), with *p.d.f.*,

$$f(x) = \begin{cases} \dfrac{1}{30}, & 0 < x < 30 \\ 0, & \text{otherwise} \end{cases}$$

The probability that he has to wait at least 20 minutes is

$$P(X \geq 20) = \int_{20}^{30} f(x)\, dx = \frac{1}{30} \int_{20}^{30} 1 \cdot dx = \frac{1}{30}(30 - 20) = \frac{1}{3}$$

**Example 8·3.** *If $X$ has a uniform distribution in [0, 1], find the distribution (p.d.f.) of $-2 \log X$. Identify the distribution also.*

**[Delhi Univ. B.Sc. (Stat. Hons.), 1989,' 86]**

**Solution.** Let $Y = -2 \log X$. Then the distribution function $G$ of $Y$ is

$$G_Y(y) = P(Y \leq y) = P(-2 \log X \leq y)$$

$$= P(\log X \geq -y/2) = P(X \geq e^{-y/2}) = 1 - P(X \leq e^{-y/2})$$

$$= 1 - \int_0^{e^{-y/2}} f(x)\, dx = 1 - \int_0^{e^{-y/2}} 1 \,.\, dx = 1 - e^{-y/2}$$

$$g_Y(y) = \frac{d}{dy} G(y) = \frac{1}{2} e^{-y/2}, \; 0 < y < \infty \qquad \ldots(*)$$

[ $\because$ as $X$ ranges in (0, 1), $Y = -2 \log X$ ranges from 0 to $\infty$ ]

**Remark.** This example illustrates that if $X \sim U[0, 1]$, then $Y = -2 \log X$, has an exponential distribution with parameter $\theta = \frac{1}{2}$. [ *c.f.* § 8·6 ] or $Y = -2 \log X$ has chi-square distribution with $n = 2$ degrees of freedom [ *c.f.* Chapter 13, § 13·2 ].

**Example 8·4.** *Show that for the rectangular distribution :*

$$f(x) = \frac{1}{2a}, -a < x < a$$

*the m.g.f. about origin is* $\frac{1}{at}$ *( sinh at ). Also show that moments of even order are given by*

$$\mu_{2n} = \frac{a^{2n}}{(2n+1)}$$

**Solution.** M.G.F. about origin is given by

$$M_X(t) = E(e^{tX}) = \int_{-a}^{a} e^{tx} f(x)\, dx = \frac{1}{2a}\int_{-a}^{a} e^{tx}\, dx$$

$$= \frac{1}{2a}\left| \frac{e^{tx}}{t} \right|_{-a}^{a} = \frac{1}{2at}(e^{at} - e^{-at}) = \frac{\sinh at}{at}$$

$$= \frac{1}{at}\left[ at + \frac{(at)^3}{3!} + \frac{(at)^5}{5!} + \ldots \right] = 1 + \frac{a^2 t^2}{3!} + \frac{a^4 t^4}{5!} + \ldots$$

Since there are no terms with odd powers of $t$ in $M(t)$, all moments of odd order about origin vanish, *i.e.*,

$$\mu'_{2n+1} \text{ (about origin)} = 0$$

In particular $\mu_1'$ (about origin) $= 0$, *i.e.*, mean $= 0$

Thus $\mu_r'$ (about origin) $= \mu_r$ (since mean is origin)

Hence $\mu_{2n+1} = 0\ ;\ n = 0, 1, 2, \ldots$

*i.e.*, all moments of odd order about mean vanish. The moments of even order are given by

$$\mu_{2n} = \text{coefficient of } \frac{t^{2n}}{(2n)!} \text{ in } M(t) = \frac{a^{2n}}{(2n+1)}$$

**Example 8·5.** *If* $X_1$ and $X_2$ *are independent rectangular variates on* [0, 1], *find the distributions of*

(*i*) $X_1/X_2$, (*ii*) $X_1 X_2$, (*iii*) $X_1 + X_2$, and (*iv*) $X_1 - X_2$

**Solution.** We are given

$$f_{X_1}(x_1) = f_{X_2}(x_2) = 1\ ;\ 0 < x_1 < 1, 0 < x_2 < 1$$

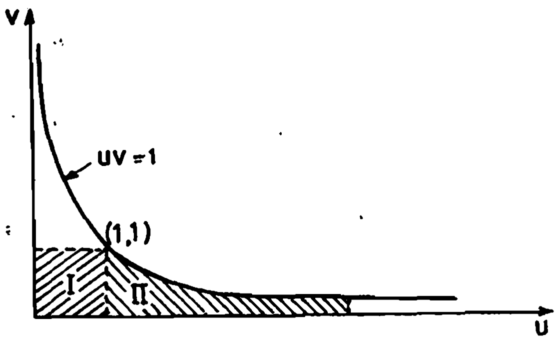
Since $X_1$ and $X_2$ are independent, their joint p.d.f. is

$$f(x_1, x_2) = f(x_1) f(x_2) = 1$$

(*i*) Let us transform to

$$u = \frac{x_1}{x_2},\ v = x_2 \quad i.e.,\quad x_1 = uv,\ x_2 = v$$

$$J = \frac{\partial(x_1, x_2)}{\partial(u, v)} = \begin{vmatrix} v & 0 \\ u & 1 \end{vmatrix} = v$$




$x_1 = 0$ maps to $u = 0,\ v = 0$

$x_1 = 1$ maps to $uv = 1$ (Rectangular hyperbola)

$x_2 = 0$ maps to $v = 0$ and $x_2 = 1$ maps to $v = 1$ .

The joint p.d.f. of $U$ and $V$ becomes

$$g(u, v) = f(x_1, x_2)\,|J| = v\ ;\ 0 < u < \infty,\ 0 < v < \infty$$

To obtain the marginal distribution of $U$, we have to integrate out $v$.

In region $(I)$,

$$g_1(u) = \int_0^1 v\,dv = \left|\frac{v^2}{2}\right|_0^1 = \tfrac{1}{2},\ 0 \le u \le 1$$

In region $(II)$,

$$g_1(u) = \int_0^{1/u} v\,dv = \left|\frac{v^2}{2}\right|_0^{1/u} = \frac{1}{2u^2},\ 1 < u < \infty$$

Hence the distribution of $U = \dfrac{X_1}{X_2}$ is given by

$$g(u) = \tfrac{1}{2},\ 0 \le u \le 1$$

$$= \frac{1}{2u^2},\ 1 \le u < \infty$$

(ii) Let $u = x_1 x_2$, $v = x_1$, i.e., $x_1 = v,\ x_2 = \dfrac{u}{v}$

$$J = \begin{vmatrix} 0 & 1 \\ \dfrac{1}{v} & -\dfrac{u}{v^2} \end{vmatrix} = -\frac{1}{v}$$

$x_1 = 0$ maps to $v = 0$, $x_1 = 1$ maps to $v = 1$

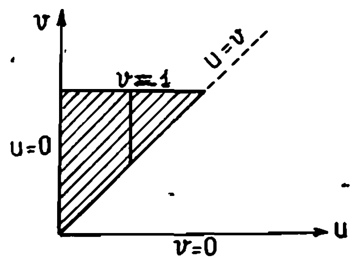
$x_2 = 0$ maps to $u = 0$, and $x_2 = 1$ maps to $u = v$

Moreover, $v = \dfrac{u}{x_2} \Rightarrow v \geq u$ (since $0 < x_2 < 1$),

The joint p.d.f. of $U$ and $V$ is

$$g(u, v) = f(x_1, x_2)\,|J| = \frac{1}{v}\ ;\ 0 < u < 1,\ 0 < v < 1$$

$$g(u) = \int_u^1 \frac{1}{v}\,dv = [\log v]_u^1 = -\log u,\ 0 < u < 1$$




(*iii*) and (*iv*). Let $u = x_1 + x_2$,
$v = x_1 - x_2$

i.e., $x_1 = \dfrac{u+v}{2}$, $x_2 = \dfrac{u-v}{2}$

$x_1 = 0 \Rightarrow u + v = 0$ i.e., $v = -u$
$x_2 = 0 \Rightarrow u - v = 0$ i.e., $v = u$
$x_1 = 1 \Rightarrow u + v = 2$
$x_2 = 1 \Rightarrow u - v = 2$

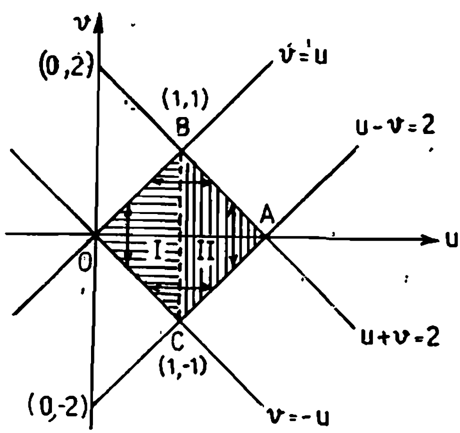
$$\text{and } J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

$\therefore\ g(u, v) = f(x_1, x_2)\,|J| = \frac{1}{2},\ 0 < u < 2,\ -1 < v < 1$

In region (*I*), (see figure below)

$$g_1(u) = \int_{-u}^{u} \frac{1}{2}\,dv = \frac{1}{2}\,|\,v\,|_{-u}^{u} = u$$

and in region (*II*),



$$g_2(u) = \int_{u-2}^{2-u} \tfrac{1}{2}\, dv = \tfrac{1}{2} \left| v \right|_{u-2}^{2-u} = 2 - u$$

$$\therefore \quad g(u) = \begin{cases} u, & 0 < u < 1 \\ 2 - u, & 1 < u < 2 \end{cases}$$

For the distribution of $V$, we split the region as: $OAB$ and $OAC$

In region $OAB$:

$$h_1(v) = \int_{v}^{2-v} \tfrac{1}{2}\, du = \tfrac{1}{2}[2 - v - v] = 1 - v, \; 0 < v < 1$$

In region $OAC$:

$$h_2(v) = \int_{-v}^{2+v} \tfrac{1}{2}\, du = \tfrac{1}{2}[2(1+v)] = 1 + v, \; -1 < v < 0$$

Hence the distribution of $V = X_1 - X_2$ is given by

$$h(v) = \begin{cases} 1 - v, & 0 < v < 1 \\ 1 + v, & -1 < v < 0 \end{cases}$$

**Example 8·6.** *If $X$ is a random variable with a continuous distribution function $F$, then $F(X)$ has a uniform distribution on [0, 1].*

**[Delhi Univ. B.Sc. (Stat. Hons.), 1992, 1987, '85]**

**Solution.** Since $F$ is a distribution function, it is non-decreasing. Let $Y = F(X)$, then the distribution function $G$ of $Y$ is given by

$$G_Y(y) = P(Y \le y) = P[F(X) \le y] = P[X \le F^{-1}(y)],$$

the inverse exists, since $F$ is non-decreasing and given to be continuous.

$$\therefore \quad G_Y(y) = F[F^{-1}(y)],$$

since $F$ is the distribution function of $X$.

$$\therefore \quad G_Y(y) = y$$

Therefore the p.d.f. of $Y = F(X)$ is given by:

$$g_Y(y) = \frac{d}{dy}[G_Y(y)] = 1$$

Since $F$ is a d.f., $Y$ takes the values in the range [0, 1].

Hence $g_Y(y) = 1, 0 \le y \le 1$

$\Rightarrow$ $Y$ is a uniform variate on $[0, 1]$.

**Remark.** Suppose $X$ is a random variable with p.d.f.,

$$f_X(x) = \begin{cases} e^{-x}, & x \ge 0 \\ 0, & \text{otherwise} \end{cases}$$

then

$$F(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1 - e^{-x}, & \text{if } x \ge 0 \end{cases}$$

Then by above result $F(X) = 1 - e^{-X}$ is uniformly distributed on [0, 1].

**Example 8·7.** *If $X$ and $Y$ are independent rectangular variates for the range $-a$ to $a$ each, then show that the sum $X + Y = U$, has the probability density*

$$\varphi(u) = \frac{2a+u}{4a^2}, \; -2a \le u \le 0$$

$$\varphi(u) = \frac{2a-u}{4a^2}, \; 0 \le u \le 2a$$

**Solution.** Since $X$ and $Y$ are independent rectangular variates, each in the interval $(-a, a)$, we have

$$f_1(x) = \begin{cases} \dfrac{1}{2a}, & -a < x < a \\ 0, & \text{elsewhere} \end{cases}$$

and
$$f_2(y) = \begin{cases} \dfrac{1}{2a}, & -a < y < a \\ 0, & \text{elsewhere} \end{cases}$$

Hence by compound probability theorem, the joint probability differential of $X$ and $Y$ is given by

$$dP(x, y) = f_1(x) f_2(y)\, dx\, dy = \frac{1}{4a^2}\, dx\, dy, \; -a < (x, y) < a$$

Let us define new variables $U$ and $V$ as follows :

$$u = x + y, \qquad v = x - y$$

$$\Rightarrow \qquad x = \frac{u+v}{2} \quad \text{and} \quad y = \frac{u-v}{2}$$

Jacobian of the transformation $J$ is given by

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ \dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

Thus the probability differential of $U$ and $V$ becomes

$$dG(u, v) = \frac{1}{4a^2} |J|\, du\, dv = \frac{1}{8a^2}\, du\, dv \qquad \ldots(*)$$

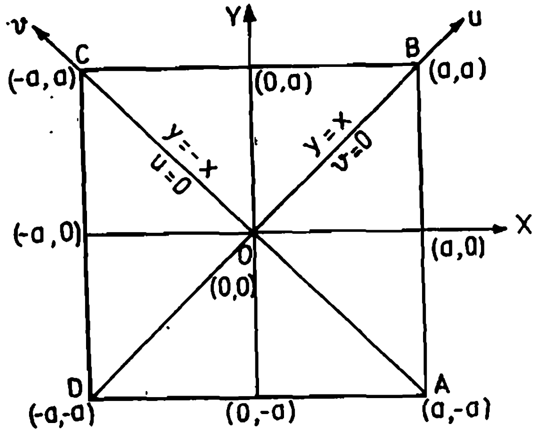
Integrating w.r.to. $v$ over specified range, we can find the distribution of $U$.

Let us consider the region to the left of $v-$ axis, *i.e.*, to the left of the line $AC$. In this region, the values of $v$ are bounded by the lines $x = -a$ and $y = -a$.

For fixed values of $u$,

$$x = -a \quad \Rightarrow \quad \frac{u+v}{2} = -a \quad \Rightarrow \quad v = -(u + 2a)$$

and
$$y = -a \quad \Rightarrow \quad \frac{u-v}{2} = -a \quad \Rightarrow \quad v = (u + 2a)$$




Thus integrating (*) w.r.to. $v$ between the limits $-(u+2a)$ and $(u+2a)$, the distribution of $U$ becomes

$$g_1(u)\,du = \int_{-(u+2a)}^{u+2a} \frac{1}{8a^2}\cdot du\,dv = \frac{1}{8a^2}\,|\,v\,|_{-(u+2a)}^{u+2a}\;du = \frac{u+2a}{4a^2}\,du$$

In the region to the left of $v$-axis, *i.e.*, below the line $AC$, $u$ varies from the points $(x=-a,\ y=-a)$ to the point $(x=0, y=0)$ and since $u=x+y$, in this region $u$ lies between $(-a-a)$ and $(0+0)$, *i.e.*, between $-2a$ to $0$.

$$\therefore \qquad g_1(u)\,du = \frac{u+2a}{4a^2}\,, \quad -2a \le u \le 0$$

In the region to the right of $v$–axis, *i.e.*, above the line $AC$, the values of $v$ are bounded by the lines $x=a$ and $y=a$ and for fixed values of $u$,

$$x=a \quad\Rightarrow\quad \frac{u+v}{2}=a \quad\Rightarrow\quad v=2a-u$$

$$y=a \quad\Rightarrow\quad \frac{u-v}{2}=a \quad\Rightarrow\quad v=-(2a-u)$$

In this region $u$ varies from the point $(x=0, y=0)$ to the point $(x=a, y=a)$, i.e., $u=x+y$ varies from 0 to $2a$. Thus integrating (*) w.r.to. $v$ between the limits $-(2a-u)$ to $(2a-u)$, we get the distribution of $U$ as

$$g_1(u)\,du = \int_{-(2a-u)}^{2a-u} \frac{1}{8a^2}\,du\,dv = \frac{1}{8a^2}\,|\,v\,|_{-(2a-u)}^{2a-u}\,du$$

$$= \frac{2a-u}{4a^2}\,du,\ 0 \le u \le 2a$$

For an alternative and simpler solution, see Remark 5 to § 8·1·5, (Triangular Distribution).

**Example. 8·8.** *On the x-axis $(n+1)$ points are taken independently between the origin and $x=1$, all positions being equally likely. Show that probability* *that the $(k+1)$ th of these points, counted from the origin, lies in the interval $x - \frac{1}{2}\,dx$ to $x + \frac{1}{2}\,dx$ is*

$$\binom{n}{k}(n+1)\,x^k\,(1-x)^{n-k}\,dx$$

*Verify that integral of this expression from $x = 0$ to $x = 1$ is unity.*

**Solution.** Here $X$ is given to be a random variable uniformly distributed on $[0, 1]$.

$$\therefore \qquad f_X(x) = 1,\ 0 \le x \le 1$$

$$\text{Now} \qquad P(0 < X < x) = \int_0^x f(x)\,dx = \int_0^x 1\,.\,dx = x \qquad \ldots(1)$$

$$\therefore \qquad P(X > x) = 1 - P(X \le x) = 1 - x \qquad \ldots(2)$$

$$\text{Also } P\left(x - \frac{dx}{2} < X < x + \frac{dx}{2}\right) = \int_{x - \frac{dx}{2}}^{x + \frac{dx}{2}} f(x)\,dx = dx \qquad \ldots(3)$$

Required probability '$p$' is given by

$p = P$ {out of $(n+1)$ points, $k$ points lie in the closed interval $\left[0, x - \frac{dx}{2}\right]$ and out of the remaining $(n+1-k)$ points, $(n-k)$ points lie in $\left[x + \frac{dx}{2}, 1\right]$ and one point lies in $\left(x - \frac{dx}{2}, x + \frac{dx}{2}\right)$}

$$= \left[\binom{n+1}{k} x^k\right] \times \left[\binom{n+1-k}{n-k}(1-x)^{n-k}\right] \times dx,$$

on using (1), (2) and (3) respectively.

$$\therefore \qquad p = \frac{(n+1)!}{k!\,(n+1-k)!}\,.\,x^k\,.\,\frac{(n+1-k)!}{(n-k)!}\,.\,(1-x)^{n-k}\,dx$$

$$= \binom{n}{k}(n+1)\,x^k\,(1-x)^{n-k}\,dx$$

To prove that the area of this expression from $x = 0$ to $x = 1$ is unity, use Beta-integral

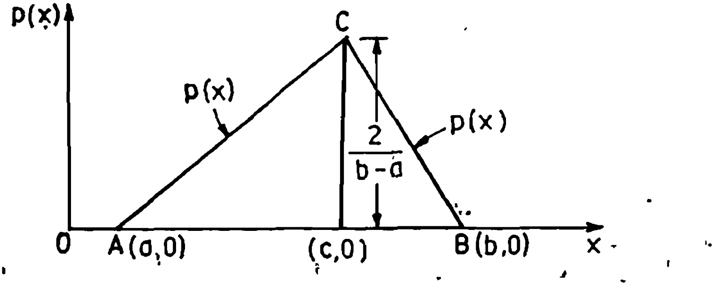
$$\int_0^1 x^{m-1}(1-x)^{n-1}\,dx = B(m, n) = \frac{\Gamma m\,\Gamma n}{\Gamma(m+n)}\,;\ m > 0, n > 0\,.$$

**8·1·5. Triangular Distribution.** A random variable $X$ is said to have a triangular distribution in the interval $(a, b)$, if its p.d.f. is given by:

$$f(x) = \begin{cases} 2(x-a)/\{(b-a)(c-a)\} & ;\ a < x \le c \\ 2(b-x)/\{(b-a)(b-c)\} & ;\ c < x < b \end{cases} \qquad \ldots(8\cdot 2a)$$

**Remarks. 1.** We write $X \sim$ Trg. $(a, b)$, with peak at $x = c$. The graph of the **p.d.f. is shown in the diagram on page 8·11.**

2. The distribution is so called because the graph of its p.d.f. is a triangle with peak at $x = c$.




3. The m.g.f. of Trg $(a, b)$ variate, with peak at $x = c$ is given by:

$$M_X(t) = \int_a^b e^{tx} f(x)\, dx = \left(\int_a^c + \int_c^b\right) e^{tx} f(x)\, dx.$$

$$= \frac{2}{(b-a)(c-a)} \int_a^c e^{tx}(x-a)\, dx + \frac{2}{(b-a)(b-c)} \int_c^b e^{tx} \cdot (b-x)\, dx$$

$$= \frac{2}{t^2}\left\{\frac{e^{at}}{(a-b)(a-c)} + \frac{e^{ct}}{(c-a)(c-b)} + \frac{e^{bt}}{(b-a)(b-c)}\right\}; a < b < c$$

(On integration by parts) ...(8·2$b$)

4. In particular, taking $a = 0$, $c = 1$ and $b = 2$, in (8·2$a$), the p.d.f. of the Trg (0, 2) variate with peak at $x = 1$ is given by:

$$f(x) = \begin{cases} x; & 0 \le x \le 1 \\ 2 - x; & 1 \le x \le 2 \\ 0, & \text{otherwise} \end{cases} \qquad ...(8\cdot 2\,c)$$

and its m.g.f. is, $M_X(t) = (e^t - 1)^2/t^2$, ...(8·2$d$)

**which is left as an exercise to the reader.**

5. In particular, replacing $a$ by $-2a$, $b$ by $2a$ and $c$ by $0$, the p.d.f. of triangular distribution on the interval $(-2a, 2a)$ with peak at $x = 0$ is given by:

$$f(x) = \begin{cases} (2a + x)/4a^2; & -2a < x < 0 \\ (2a - x)/4a^2; & 0 < x < 2a \end{cases} \qquad ...(8\cdot 2e)$$

The m.g.f. of (8·2$e$) is given by :

$$M_X(t) = \int_{-2a}^{2a} e^{tx} f(x)\, dx$$

$$= \frac{1}{4a^2}\left[\int_{-2a}^{0} e^{tx} \cdot (2a + x)\, dx + \int_0^{2a} e^{tx}(2a - x)\, dx\right]$$

$$= \frac{1}{4a^2}\left[ e^{tx}\left\{ \frac{2a+x}{t} - \frac{1}{t^2} \right\} \right]_{-2a}^{0} + \frac{1}{4a^2}\left[ e^{tx}\left\{ \frac{2a-x}{t} + \frac{1}{t^2} \right\} \right]_{0}^{2a}$$

[On integrating by parts]

$$= \frac{1}{4a^2}\left[ -\frac{2}{t^2} + \frac{1}{t^2}\left\{ e^{2at} + e^{-2at} \right\} \right]$$

$$= \frac{1}{4a^2 t^2}\left\{ e^{2at} + e^{-2at} - 2 \right\} = \left[ \frac{e^{at} - e^{-at}}{2\,at} \right]^2 \qquad ...(8{\cdot}2f)$$

**Aliter.** We may obtain (8·2 *f*) directly from (8·2 *b*) on replacing a by $-2a$, $b$ by $2a$ and $c$ by 0.

**Example 8·9.** *If $X$ and $Y$ are i.i.d. $U[-a, a]$ variates, find the p.d.f. of $Z = X + Y$ and identify the distribution.*

**Solution.** Since $X$ and $Y$ are i.i.d. $U[-a, a]$, we have : [*c.f.* § 8·1·2.],

$$M_X(t) = M_Y(t) = (e^{at} - e^{-at})/(2\,at) \qquad ...(*)$$

$$M_{X+Y}(t) = M_X(t)\, M_Y(t) = \left[ \frac{e^{at} - e^{-at}}{2at} \right]^2, \qquad ...(**)$$

since $X$ and $Y$ are independent.

But (**) is the m.g.f. of Trg $(-2a, 2a)$ variate with peak at $x = 0$

[*c.f.* Remark 5, equation (8·2*f*) ]

Hence by uniqueness theorem of m.g.f., $Z = X + Y \sim$ Trg $(-2a, 2a)$ with p.d.f. as given in (8·2 *e*), Remark 5.

**Aliter** $\quad M_{X+Y}(t) = \dfrac{1}{4a^2 t^2}\left[ e^{2at} - 2 + e^{-2at} \right]$ [From (**)]

$$= \frac{2}{t^2}\left[ \frac{e^{-2at}}{(-2a-0)(-2a-2a)} + \frac{e^{0.t}}{(0+2a)(0-2a)} + \frac{e^{2at}}{(2a-0)(2a+2a)} \right]$$

which is of the form (8·2 *b*), [ *c.f.* Remark 3 ], with $a$ replaced by $-2a$ and $b$ replaced by $2a$ and $c$ by 0 . Hence $X + Y \sim$Trg $(-2a, 2a)$ with p.d.f. $p(x)$ given in (8·2*e*) .

**Remarks 1.** The distribution of $X + Y$ has also been obtained in Example 8·7.

**2.** Similarly we can find the distribution of $X - Y$.

$$M_{X-Y}(t) = M_X(t)\,.\,M_Y(-t) = \left[ \frac{e^{at} - e^{-at}}{2at} \right]^2 \qquad \text{[From (*)]}$$

$\Rightarrow X - Y \sim$Trg $(-2a, 2a)$, with peak at $x = 0$ .

## EXERCISE 8 (a)

**1.** The bus company $A$ schedules a north bound bus every 30 minutes at a certain bus-stop. A man comes to the stop at a random time. Let the random variable $X$ count the number of minutes he has to wait for the next bus. Assume $X$ has a uniform distribution over the interval (0, 30). This is how we interpret the statement that he enters the station at the random time].

(*i*) For each $k = 5, 10, 15, 20, 30$ compute the probability that he has to wait at least $k$ minutes for the next bus.

(*ii*) A competitor, the bus company $B$ is allowed to schedule a north bound bus every 30 minutes at the same station but at least 5 minutes must elapse between the arrivals of the competitive buses. Assume the passengers come at the bus stop at random times and always board the first bus that arrives. Show that the company $B$ can arrange its schedule so that it receives five times as many passengers as that of its competitor.

**2.** (*a*) A random variable $X$ has a uniform distribution over $(-3, 3)$, compute

(*i*) $P(X = 2)$, $P(X < 2)$, $P(|X| < 2)$ and $P(|X - 2| < 2)$

(*ii*) Find $k$ for which $P(X > k) = 1/3$. **[Gorakhpur Univ. B.Sc. 1992)**

(*b*) Suppose that $X$ is uniformly distributed over $(-\alpha, +\alpha)$, where $\alpha > 0$. Determine $\alpha$ so that

(*i*) $P(X > 1) = 1/3$, (*ii*) $P(X < 1/2) = 0{\cdot}3$ and

(*ii*) $P(|X| < 1) = P(|X| > 1)$.

**Ans.** (*i*) $\alpha = 3$, (*ii*) $\alpha = 5/6$, (*iii*) $\alpha = 2$.

(*c*) Calculate the coefficient of variation for the rectangular distribution in $(0, b)$ given that the probability law of the distribution is

$$P(X \le t) = \frac{t}{b}$$

(*d*) If $X$ is uniformly distributed over $[1, 2]$, find $z$ so that

$$P(X > z + \mu_x) = \frac{1}{4} \qquad \textbf{(Ans. } Z = \frac{1}{4}\textbf{)}.$$

**3** (*a*). If a random variable $X$ has the density function $f(x)$, prove that

$$Y = \int_{-\infty}^{x} f(x)\, dx$$

has a rectangular distribution over (0, 1). If

$$f(x) = \tfrac{1}{2}(x - 1), \quad 1 \le x \le 3$$
$$= 0, \text{ otherwise}$$

determine what interval for $Y$ will correspond to the interval

$$1{\cdot}1 \le X \le 2{\cdot}9.$$

**Ans.** $y = F(x) = (x - 1)^2/4$ ; $1 \le x \le 3$ ; $0{\cdot}0025 \le y \le 0{\cdot}9025$

(*b*). Show that whatever be the distribution function $F(x)$ of a r.v. $X$,

$$P[a \le F(X) \le b] = b - a, \quad 0 \le (a, b) \le 1.$$

**[Delhi Univ. B.Sc. (Stat. Hons), 1986]**

**Hint.** $Y = F(X) \sim U[0, 1]$.

**4.** (*a*) For the rectantular distribution,

$$f(x) = \frac{1}{2a}\ ; -a \le x \le a\ .$$
$$= 0, \text{ otherwise},$$

show that the moments of odd order are zero, and $\mu_{2r} = a^{2r}/(2r+1)$.

**[Madurai Kamraj Univ. B.Sc. 1992]**

(*b*) A distribution is given by

$$f(x)\,dx = \frac{1}{2a}\,dx,\ -a \le x \le a$$

Find the first four central moments and obtain $\beta_1$ and $\beta_2$.

**[Delhi Univ B.Sc. Oct., 1992; Madras Univ. B.Sc., 1991]**

(*c*) For a rectangular distribution

$$dP = k\,dx,\ 1 \le x \le 2,$$

show that Arithmetic mean > Geometric mean > Harmonic mean.

**[Vikram Univ. B.Sc. 1993)**

(*d*) If the random variable $X$ follows the rectangular distribution with p.d.f.,

$$f(x) = 1/\theta,\ 0 \le x \le \theta\ ,$$

derive the first four moments and the skewness and kurtosis conffīcients of the distribution.

(*e*) Let $X$ and $Y$ be independent variates which are uniformly distributed over the unit interval (0,1). Find the distribution function and the p.d.f. of random variable $Z = X + Y$. Is $Z$ a uniformly distributed variable ? Give reasons.

**[Delhi Univ. B.Sc. (Maths. Hons.), 1986]**

5. Let $X_1$ and $X_2$ be independent random variables unifromly distributed over the interval (0, 1). Find

(*i*) $P(X_1 + X_2 < 0{\cdot}5)$, (*ii*) $P(X_1 - X_2 < 0{\cdot}5)$,

(*iii*) $P(X_1^2 + X_2^2 < 0{\cdot}5)$, (*iv*) $P(e^{-X_1} < 0{\cdot}5)$, and (*v*) $P(\cos \pi X_2 < 0{\cdot}5)$.

**Ans.** (*i*) 0·125, (*ii*) 0·875. (*iii*) 0·393, (*iv*) $1 - \log 2$, and (*v*) $2/3$.

6. A random variable $X$ is uniformly distributed over (0, 1), find the probability density functions of

(*i*) $Y = X^2 + 1$, and (*ii*) $Z = 1/(X+1)$.

7. (*a*) If the random variable $X$ is uniformly distributed over $(0, \frac{1}{2}\pi)$, compute the expectation of the function $\sin X$. Also find the distribution of $Y = \sin X$, and show that the mean of this distribution is the same as the above expectation.

**Ans.** $2/\pi$, $f_Y(y) = 2/(\pi\sqrt{1-y^2}),\ 0 < y < 1$.

(*b*) If $X \sim U[-\pi/2, \pi/2]$ distributed, find the p.d.f. of $Y = \tan X$.

**[Delhi Univ. B.A. Hons. (Spl. Course-Statistics), 1989]**

8. (*a*) Show that for the rectangular distribution :

$$dF = dx,\ 0 \le x < 1$$

$\mu'_1$ (about origin) = 1/2, variance = 1/12 and mean deviation about mean = 1/4. **[Madras Univ B.Sc. Sept. 1991; Delhi U. B.Sc. Sept. 1992]**

(*b*) Find the characteristic function of the random variable $Y = \log F(X)$ where $F(X)$ is the distribution function of a random variable $X$. Evaluate the $r$th moment of $Y$.

9. If $X \sim U[0, 1]$, find the distribution of $Y = 1/X$. Find $E(1/X)$, if it exists.

**Ans.** $g_Y(y) = 1/y^2$; $1 \le y < \infty$; $E(Y) = E(1/X)$ does not exist.

10. Let $X$ be uniformly distributed on $[-1, 1]$. Find the distribution function and hence the p.d.f. of $Y = X^2$. **[Delhi Univ. B.Sc. (Maths. Hons.), 1988]**

11. Let $f_X(x) = 6x(1-x)$; $0 \le x \le 1$. Find $y$ as a function of $x$ such that $Y$ has p.d.f.

$$g(y) = 3(1-\sqrt{y})\,; \quad 0 \le y \le 1$$

**[Delhi Univ. B.A. Hons. (Spl. Course-Statistics), 1988]**

**Hint.** $$F(x) = \int_0^x f(x)\,dx = 3x^2 - 2x^3 \sim U[0, 1]$$

$$G(y) = \int_0^y g(y)\,dy = 3y - 2y^{3/2} \sim U[0, 1]$$

Setting $F(x) = G(y)$, we get $y = x^2$.

12. The variates $a$ and $b$ are independently and uniformly distributed in the intervals [0,6] and [0, 9] respectively. Find the probability that $x^2 - ax + b = 0$ has two real roots.

**Ans.** $$P(b \le a^2/4) = \int_{a=0}^{1} \int_{b=0}^{a^2/4} \frac{1}{6 \times 9}\, da\, db = 1/3.$$

13. Find the probability that the roots of the equation $x^2 + 2bx + c = 0$ should be real, given that $b \sim U[-\alpha, \alpha]$ and $c \sim U[-\beta, \beta]$ are independent.

**Ans.** Probability $= P(b^2 \ge c) = 1 - P(b^2 \le c) = 1 - P(|b| \le \sqrt{c})$

$$= \int_{-\beta}^{\beta} \left( \int_{-\sqrt{c}}^{\sqrt{c}} \left(\frac{1}{2\alpha}\right)\left(\frac{1}{2\beta}\right) db \right) dc$$

14. If $a, b, c$ are randomly chosen between 0 and 1, find the probability that the quadratic equation $ax^2 + bx + c = 0$ has real roots.

**Ans.** Probability $$= P(b^2 \ge 4ac) = 1 - \int_0^1 \int_0 \int_{b=0}^{\sqrt{4ac}} 1 \; da\, dc\, db = \frac{8}{9}$$

15. (*a*) Suppose $X$ has a rectangular distribution on $(-1, 1)$. Compute $P\left[\dfrac{|X - E(X)|}{\sigma_X} \ge 2\right]$ and compare it with the upper bound given by Chebyshev's inequality.

(*b*) Compare the upper bound of the probability,

$$P\{|X - E(X)| \geq 2\sqrt{V(X)}\},$$

obtained from Chebyshev's inequality with exact probability if $X$ is uniformly distributed over $(-1, 3)$.

**Ans.** (*b*) Probability $\leq 1/4$, Exact Probability $= 0$

**16.** Two independent variates are each uniformly distributed within the range $-a$ to $+a$. Show that their sum $X$ has a probability density given by

$$f(x) = \frac{2a+x}{4a^2}, \quad -2a \leq x \leq 0$$

$$= \frac{2a-x}{4a^2}, \quad 0 \leq x \leq 2a$$

Verify that the m.g.f. calculated from the value of $f(x)$ is equal to

$$\left(\frac{1}{at} \sinh at\right)^2$$

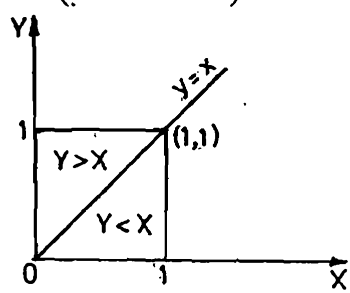
**17.** The random variables $X$ and $Y$ are independent and both have the uniform distribution on $[0, 1]$. Let $Z = |X - Y|$. Prove that, for real $\theta$,

$$\varphi(Z, \theta) = 2\left[1 + i\theta - e^{i\theta}\right]/2.$$

Hence deduce the general expression for $E(Z^n)$.

**Hint.** $$\varphi(\theta; |X - Y|) = \int_0^1 \int_0^1 e^{i\theta|x-y|} f(x, y)\, dx\, dy$$

$$= 2\int_0^1 \left(\int_0^x e^{i\theta(x-y)}\, dy\right) dx$$




**Ans.** $2/[(n+1)(n+2)]$

**18.** If $X$ and $Y$ are independently and uniformly distributed random variables in the interval $(0, 1)$, show that the distribution of $X + Y$ is given by the density function

$$f(z) = \begin{cases} z & 0 \leq z < 1 \\ 2 - z & 1 \leq z \leq 2 \\ 0 & \text{elsewhere} \end{cases}$$

[Hint. See Triangular distribution]

**19.** Ship $A$ makes radio signals to the base and the probability of the interval between consecutive signals is uniformly distributed between 4 hours and 24 hours and is zero outside this range. Ship $B$ makes radio signals to the base and the probability of the interval between consecutive signals is uniformly distributed between 10 hours and 15 hours and is zero outside this range.

(*i*) Ship $A$ has just signalled. What is the probability that it will make two further signals in the next 12 hours ?

(*ii*) Ships $A$ and $B$ have just signalled at the same time. What is the probability that Ship $A$ will make at least two further signals before ship $B$ next signals? **[Institute of Actuaries (London), April 1978]**

**20.** If $X \sim U[0, 1]$, prove that for $b < c$ fixed, $Y = (c-b)X + b$ is uniform on $[b, c]$.

**8·2. Normal Distribution.** The normal distribution was first discovered in 1733 by English mathematician De-Moivre, who obtained this continuous distribution as a limiting case of the binomial distribution and applied it to problems arising in the game of chance. It was also known to Laplace, no later than 1774 but through a historical error it was credited to Gauss, who first made reference to it in the beginning of 19th century (1809), as the distribution of errors in Astronomy. Gauss used the normal curve to describe the theory of accidental errors of measurements involved in the calculation of orbits of heavenly bodies. Throughout the eighteenth and nineteeth centuries, various efforts were made to establish the normal model as the underlying law ruling all continuous random variables. Thus, the name "*normal*" . These efforts, however, failed because of false premises. The normal model has, nevertheless, become the most important probability model in statistical analysis.

**Definition.** *A random variable X is said to ahve a normal distribution with parameters* $\mu$ *(called "mean") and* $\sigma^2$ *(called "variance") if its density function is given by the probability law :*

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left\{\frac{x-\mu}{\sigma}\right\}^2\right]$$

or $$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2},$$

$$-\infty < x < \infty, \; -\infty < \mu < \infty, \sigma > 0 \quad ...(8{\cdot}3)$$

**Remarks. 1.** A random variable $X$ with mean $\mu$ and variance $\sigma^2$ and following the normal law (8·3) is expressed by $X \sim N(\mu, \sigma^2)$

**2.** If $X \sim N(\mu, \sigma^2)$, then $Z = \dfrac{X-\mu}{\sigma}$, is a standard normal variate with

$$E(Z) = 0 \text{ and } \mathrm{Var}(Z) = 1$$

and we write $Z \sim N(0,1)$.

**3.** The p.d.f. of standard normal variate $Z$ is given by

$$\varphi(z) = \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2}, \; -\infty < z < \infty$$

and the corresponding distribution function, denoted by $\Phi(z)$ is given by

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \varphi(u)\, du$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2}\, du$$

We shall prove below two important results on the distribution function $\Phi(\cdot)$ of standard normal variate.

**Result 1.** $\Phi(-z) = 1 - \Phi(z)$

**Proof.** $\Phi(-z) = P(Z \leq -z) = P(Z \geq z)$ (By symmetry)

$$= 1 - P(Z \leq z)$$
$$= 1 - \Phi(z)$$

**Result 2.** $P(a \leq X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$, where $X \sim N(\mu, \sigma^2)$

***Proof.*** $P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right); \quad \left(Z = \frac{X-\mu}{\sigma}\right)$

$$= P\left(Z \leq \frac{b-\mu}{\sigma}\right) - P\left(Z \leq \frac{a-\mu}{\sigma}\right)$$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

**4.** The graph of $f(x)$ is a famous *'bell-shaped'* curve. The top of the bell is directly above the mean $\mu$. For large values of $\sigma$, the curve tends to flatten out and for small values of $\sigma$, it has a sharp peak.

**8·2·1. Normal Distribution as a Limiting form of Binomial Distribution.** Normal distribution is another limiting form of the binomial distribution under the following conditions :

(*i*) $n$, the number of trials is indefinitely large, *i.e.*, $n \to \infty$ and

(*ii*) neither $p$ nor $q$ is very small.

The probability function of the binomial distribution with parameters $n$ and $p$ is given by

$$p(x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!\,(n-x)!}\, p^x q^{n-x}\,;\; x = 0, 1, 2, \ldots, n \qquad \ldots(*)$$

Let us now consider the standard binomial variate :

$$Z = \frac{X - E(X)}{\sqrt{V(X)}} = \frac{X - np}{\sqrt{npq}}\,;\; X = 0, 1, 2, \ldots, n \qquad \ldots(**)$$

When $\quad X = 0,\; Z = \frac{-np}{\sqrt{npq}} = -\sqrt{np/q}$

and when $X = n$, $Z = \dfrac{n - np}{\sqrt{npq}} = \sqrt{nq/p}$

Thus in the limit as $n \to \infty$, $Z$ takes the values from $-\infty$ to $\infty$. Hence the distribution of $X$ will be a continuous distribution over the range $-\infty$ to $\infty$.

We want the limiting form of (*) under the above two conditions. Using Stirling's approximation to $r!$ for large $r$, *viz.*,

$$\lim_{r \to \infty} r! \simeq \sqrt{2\pi}\, e^{-r}\, r^{r+(1/2)},$$

we have in the limit as $n \to \infty$ and consequently $x \to \infty$,

$$\lim p(x) = \lim \left[ \frac{\sqrt{2\pi}\, e^{-n}\, n^{n+\frac{1}{2}}\, p^x\, q^{n-x}}{\sqrt{2\pi}\, e^{-x}\, x^{x+\frac{1}{2}}\, \sqrt{2\pi}\, e^{-(n-x)}\, (n-x)^{n-x+\frac{1}{2}}} \right]$$

$$= \lim \left[ \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{npq}} \cdot \frac{(np)^{x+\frac{1}{2}}\, (nq)^{n-x+\frac{1}{2}}}{x^{x+\frac{1}{2}}\, (n-x)^{n-x+\frac{1}{2}}} \right]$$

$$= \lim \left[ \frac{1}{\sqrt{2\pi}\, \sqrt{npq}} \left( \frac{np}{x} \right)^{x+\frac{1}{2}} \left( \frac{nq}{n-x} \right)^{n-x+\frac{1}{2}} \right] \qquad \text{...(***)}$$

From (**), we have

$$X = np + Z\sqrt{npq} \quad \Rightarrow \quad \frac{X}{np} = 1 + Z\sqrt{q/(np)}$$

Also

$$n - X = n - np - Z\sqrt{npq} = nq - Z\sqrt{npq}$$

$$\therefore \quad \frac{n - X}{nq} = 1 - Z\sqrt{p/(nq)}\ . \quad \text{Also } dz = \frac{1}{\sqrt{npq}}\, dx$$

Hence the probability differential of the distribution of $Z$, in the limit is given from (***) by

$$d\,G(z) = g(z)\, dz = \lim_{n \to \infty} \left[ \frac{1}{\sqrt{2\pi}} \times \frac{1}{N} \right] dz \qquad \text{...(8·4)}$$

where $N = \left[ \dfrac{x}{np} \right]^{x+\frac{1}{2}} \left[ \dfrac{n-x}{nq} \right]^{n-x+\frac{1}{2}}$

$$\log N = (x + \tfrac{1}{2}) \log(x/np) + (n - x + \tfrac{1}{2}) \log\{(n-x)/nq\}$$

$$= (np + z\sqrt{npq} + \tfrac{1}{2}) \log[1 + z\sqrt{(q/np)}]$$

$$+ (nq - z\sqrt{npq} + \tfrac{1}{2}) \log[1 - z\sqrt{(p/nq)}]$$

$$= (np + z\sqrt{npq} + \tfrac{1}{2}) [z \cdot \sqrt{(q/np)} - \tfrac{1}{2} z^2 (q/np) + \tfrac{1}{3} z^3 (q/np)^{3/2} - \dots]$$

$$+ (nq - z\sqrt{npq} + \tfrac{1}{2}) \left[ - z\sqrt{(p/nq)} - \tfrac{1}{2} z^2 (p/nq) - \tfrac{1}{3} z^3 (p/nq)^{3/2} - \ldots \right]$$

$$= \Bigg[ \Bigg\{ z\sqrt{npq} - \tfrac{1}{2} qz^2 + \tfrac{1}{3} z^3 \frac{q^{3/2}}{\sqrt{np}} + z^2 q - \tfrac{1}{2} z^3 \frac{q^{3/2}}{\sqrt{np}} + \tfrac{1}{2} z\sqrt{q/np} - \tfrac{1}{4} z^2 \frac{q}{np} + \ldots \Bigg\}$$

$$+ \Bigg\{ - z\sqrt{npq} - \tfrac{1}{2} z^2 p - \tfrac{1}{3} z^3 \frac{p^{3/2}}{\sqrt{nq}} + z^2 p + \tfrac{1}{2} z^3 \frac{p^{3/2}}{\sqrt{np}} - \tfrac{1}{2} z\sqrt{p/nq} - \tfrac{1}{4} z^2 \frac{p}{np} + \ldots \Bigg\} \Bigg]$$

*i.e.*,

$$\log N = \left[ -\frac{1}{2} z^2 (p+q) + z^2 (p+q) + \frac{z}{2\sqrt{n}} \left\{ \sqrt{q/p} + \sqrt{p/q} \right\} + 0 \left\{ n^{-1/2} \right\} \right]$$

$$= \frac{z^2}{2} + 0\,(n^{-1/2}) \rightarrow \frac{z^2}{2} \text{ as } n \rightarrow \infty$$

$$\therefore \quad \lim_{n \to \infty} \log N = \frac{z^2}{2} \quad \Rightarrow \quad \lim_{n \to \infty} N = e^{z^2/2}$$

Substituting in (8·4), we get

$$d\,G(z) = g(z)\,dz = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz, \; -\infty < z < \infty \qquad \ldots (8\cdot4\,a)$$

Hence the probability function of $Z$ is

$$g(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \; -\infty < z < \infty \qquad \ldots (8\cdot4\,b)$$

This is the probability density function of the *normal distribution* with mean 0 and unit variance.

If $X$ is normal variate with mean $\mu$ and s.d. $\sigma$ then $Z = (X - \mu)/\sigma$ is standard normal variate. Jacobian of transformation is $1/\sigma$. Hence substituting in {8·4 ($b$)}, the p.d.f. of a normal variate $X$ with $E(X) = \mu$, Var $(X) = \sigma^2$ is given by

$$f_X(x) = \begin{cases} \dfrac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, & -\infty < x < \infty \\ 0, \text{ otherwise} & \end{cases}$$

**Remark.** Normal distribution can also be obtained as a limiting case of Poisson Distribution with the parameter $\lambda \rightarrow \infty$.

**8·2·2. Chief Characteristics of the Normal Distribution and Normal Probability Curve.** The normal probability curve with mean $\mu$ and standard deviation $\sigma$ is given by the equation

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \; -\infty < x < \infty$$

and has the following properties :

(*i*) The curve is bell shaped and symmetrical about the line $x = \mu$.

(*ii*) Mean, median and mode of the distribution coincide.

(*iii*) As $x$ increases numerically, $f(x)$ decreases rapidly, the maximum probability occurring at the point $x = \mu$, and given by $[p(x)]_{max} = \dfrac{1}{\sigma\sqrt{2\pi}}$.

(*iv*) $\beta_1 = 0$ and $\beta_2 = 3$.

(*v*) $\mu_{2r+1} = 0, (r = 0, 1, 2, \ldots)$,
and $\mu_{2r} = 1.3.5 \ldots (2r-1)\sigma^{2r}, (r = 0, 1, 2, \ldots)$.

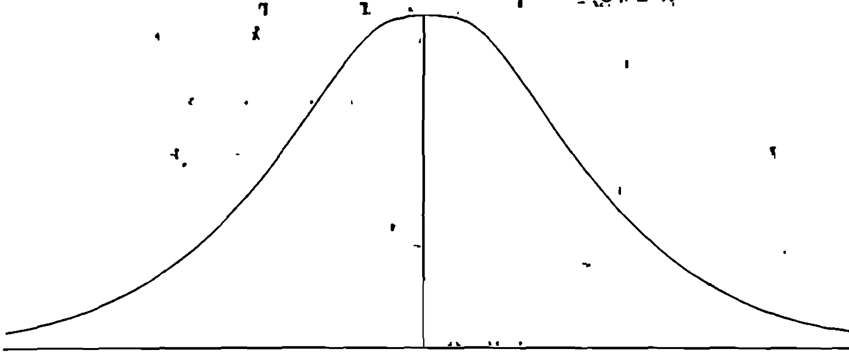
(*vi*) Since $f(x)$ being the probability, can never be negative, no portion of the curve lies below the $x$-axis.

(*vii*) Linear combination of independent normal variates is also a normal variate.

(*viii*) $x$-axis is an asymptote to the curve.

(*ix*) The points of inflexion of the curve are given by

$$\left[x = \mu \pm \sigma, f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2}\right]$$




(*Normal Probability Cuve*)

(*x*) Mean deviation about mean is $\sqrt{2/\pi}\,\sigma \approx \frac{4}{5}\sigma$ (approx.) $\quad$ Q.D. $= \dfrac{Q_3 - Q_1}{2} \simeq \dfrac{2}{3}\sigma$

We have (approximately)

$$Q.D. : M.D. : S.D. :: \frac{2}{3}\sigma : \frac{4}{5}\sigma : \sigma :: \frac{2}{3} : \frac{4}{5} : 1$$

$$\Rightarrow \quad Q.D. : M.D. : S.D. :: 10 : 12 : 15$$

(*xi*) *Area Property*

$$P(\mu - \sigma < X < \mu + \sigma) = 0{\cdot}6826$$
$$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0{\cdot}9544$$
$$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0{\cdot}9973$$

The following table gives the area under the normal probability curve for some important values of standard normal variate $Z$.

| *Distances from the mean ordinates in terms of* $\pm\sigma$ | *Area under the curve* |
|---|---|
| $Z = \pm 0{\cdot}745$ | $50\% = 0{\cdot}50$ |
| $Z = \pm 1{\cdot}00$ | $68{\cdot}26\% = 0{\cdot}6826$ |
| $Z = \pm 1{\cdot}96$ | $95\% = 0{\cdot}95$ |
| $Z = \pm 2{\cdot}0$ | $95{\cdot}44\% = 0{\cdot}9544$ |
| $Z = \pm 2{\cdot}58$ | $99\% = 0{\cdot}99$ |
| $Z = \pm 3{\cdot}0$ | $99{\cdot}73\% = 0{\cdot}9973$ |

(*xii*) If $X$ and $Y$ are independent standard normal variates, then it can be easily proved that $U = X + Y$ and $V = X - Y$ are independently distributed, $U \sim N(0, 2)$ and $V \sim N(0, 2)$.

We state (without proof) the converse of this result which is due to D. Bernstein.

**Bernstein's Theorem.** If $X$ and $Y$ are independent and identically distributed random variables with finite varaince and if $U = X + Y$ and $V = X - Y$ are independent, then all r.v.'s $X$, $Y$, $U$ and $V$ are normally distributed.

(*xiii*) We state below another result which characterises the normal distribution.

If $X_1, X_2, \ldots, X_n$ are *i.i.d. r.v.'s* with finite variance, then the common distribution is normal if and only if :

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\text{or} \qquad \sum_{i=1}^{n} X_i \text{ and } \sum_{i=1}^{n}(X_i - \bar{X})^2$$

are independent. **[For 'If part', see Theorem 13.5]**

In the following sequences we shall establish some of these properties.

**8·2·3. Mode of Normal Distribution.** Mode is the value of $x$ for which $f(x)$ is maximum, *i.e.*, mode is the solution of

$$f'(x) = 0 \text{ and } f''(x) < 0$$

For normal distribution with mean $\mu$ and standard deviation $\sigma$,

$$\log f(x) = c - \frac{1}{2\sigma^2}(x - \mu)^2,$$

where $c = \log(1/\sqrt{2\pi}\,\sigma)$, is a constant.

Differentiating w.r.t. $x$, we get

$$\frac{1}{f(x)} \cdot f'(x) = -\frac{1}{\sigma^2}(x - \mu) \Rightarrow f'(x) = -\frac{1}{\sigma^2}(x - \mu) f(x)$$

and $$f''(x) = -\frac{1}{\sigma^2}\left[1 . f(x) + (x-\mu) f'(x)\right] = -\frac{f(x)}{\sigma^2}\left[1 - \frac{(x-\mu)^2}{\sigma^2}\right] \quad ...(8·6)$$

Now $f'(x) = 0 \Rightarrow x - \mu = 0$ *i.e.*, $x = \mu$

At the point $x = \mu$, we have from (8·6)

$$f''(x) = -\frac{1}{\sigma^2}[f(x)]_{x=\mu} = -\frac{1}{\sigma^2} \cdot \frac{1}{\sigma\sqrt{2\pi}} < 0$$

Hence $x = \mu$, is the mode of the normal distribution.

**8·2·4. Median of Normal Distribution.** If $M$ is the median of the normal distribution, we have

$$\int_{-\infty}^{M} f(x)\,dx = \frac{1}{2} \Rightarrow \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{M} \exp\left\{-(x-\mu)^2/2\sigma^2\right\} dx = \frac{1}{2}$$

$$\Rightarrow \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\mu} \exp\left\{-(x-\mu)^2/2\sigma^2\right\} dx$$

$$+ \frac{1}{\sigma\sqrt{2\pi}}\int_{\mu}^{M} \exp\left\{-(x-\mu)^2/(2\sigma^2)\right\} dx = \frac{1}{2} \quad ...(8·7)$$

But $$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\mu} \exp\left\{-(x-\mu)^2/2\sigma^2\right\} dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} \exp(-z^2/2)\,dz = \frac{1}{2}$$

$\therefore$ From (8·7), we get

$$\frac{1}{2} + \frac{1}{\sigma\sqrt{2\pi}}\int_{\mu}^{M} \exp\left\{-(x-\mu)^2/2\sigma^2\right\} dx = \frac{1}{2}$$

$$\Rightarrow \quad \frac{1}{\sigma\sqrt{2\pi}}\int_{\mu}^{M} \exp\left\{-(x-\mu)^2/2\sigma^2\right\} dx = 0 \quad \Rightarrow \quad \mu = M$$

Hence for the normal distribution, Mean = Median.

**Remark.** From § 8·2·3 and § 8·2·4, we find that for the normal distribution mean, median and mode coincide. Hence the distribution is *symmetrical.*

**8·2·5. M.G.F. of Normal Distribution.** The m.g.f. (about origin) is given by

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tx} \exp\left\{-(x-\mu)^2/2\sigma^2\right\} dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left\{t(\mu+\sigma z)\right\} \exp(-z^2/2)\,dz, \qquad \left[z = \frac{x-\mu}{\sigma}\right]$$

$$= e^{\mu t}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left\{-\tfrac{1}{2}(z^2 - 2t\sigma z)\right\} dz$$

$$= e^{\mu t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\tfrac{1}{2}\left\{(z-\sigma t)^2 - \sigma^2 t^2\right\}\right] dz$$

$$= e^{\mu t + t^2\sigma^2/2} \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\tfrac{1}{2}(z-\sigma t)^2\right\} dz$$

$$= e^{\mu t + t^2\sigma^2/2} \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-u^2/2)\, du$$

Hence $$M_X(t) = e^{\mu t + t^2\sigma^2/2} \qquad \text{...(8·8)}$$

**Remark.** *M.G.F. of Standard Normal Variate.* If $X \sim N(\mu, \sigma^2)$, then standard normal variate is given by

$$Z = (X-\mu)/\sigma$$

Now $M_Z(t) = e^{-\mu t/\sigma}\, M_X(t/\sigma) = \exp(-\mu t/\sigma) \cdot \exp\left(\frac{\mu t}{\sigma} + \frac{t^2}{\sigma^2} \cdot \frac{\sigma^2}{2}\right)$

$$= \exp(t^2/2) \qquad \text{...(8·8 a)}$$

**8·2·6. Cumulant Generating Function (c.g.f.) of Normal Distribution.** The c.g.f. of normal distribution is given by

$$K_X(t) = \log_e M_X(t) = \log_e\left(e^{\mu t + t^2\sigma^2/2}\right) = \mu t + \frac{t^2\sigma^2}{2}$$

$\therefore$ Mean $= \kappa_1 =$ Coefficient of $t$ in $K_X(t) = \mu$

Variance $= \kappa_2 =$ Coefficient of $\frac{t^2}{2!}$ in $K_X(t) = \sigma^2$

and $\kappa_r =$ Coefficient of $\frac{t^r}{r!}$ in $K_X(t) = 0$ ; $r = 3, 4 \ldots$

Thus $\mu_3 = \kappa_3 = 0$ and $\mu_4 = \kappa_4 + 3\kappa_2^2 = 3\sigma^4$

Hence $$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0 \quad \text{and} \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = 3 \qquad \text{...(8·9)}$$

**8·2·7. Moments of Normal Distribution.** Odd order moments about mean are given by

$$\mu_{2n+1} = \int_{-\infty}^{\infty} (x-\mu)^{2n+1} f(x)\, dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^{2n+1} \exp\left\{-(x-\mu)^2/2\sigma^2\right\} dx$$

$$\therefore \quad \mu_{2n+1} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z)^{2n+1} \exp(-z^2/2)\, dz \qquad \left[z = \frac{x-\mu}{\sigma}\right]$$

$$= \frac{\sigma^{2n+1}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n+1} \exp(-z^2/2)\, dz = 0, \qquad \ldots(8\cdot10)$$

since the integrand $z^{2n+1}\, e^{-z^2/2}$ is an odd function of $z$.

Even order moments about mean are given by

$$\mu_{2n} = \int_{-\infty}^{\infty} (x-\mu)^{2n} f(x)\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z)^{2n} \exp(-z^2/2)\, dz$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n} \exp(-z^2/2)\, dz$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}} \cdot 2 \int_{0}^{\infty} z^{2n} \exp(-z^2/2)\, dz$$

(since integrand is an even function of $z$)

$$\therefore \qquad \mu_{2n} = \frac{2\sigma^{2n}}{\sqrt{2\pi}} \int_0^{\infty} (2t)^n e^{-t} \frac{dt}{\sqrt{2t}} \qquad \left[ \frac{z^2}{2} = t \right]$$

$$= \frac{2^n \cdot \sigma^{2n}}{\sqrt{\pi}} \int_0^{\infty} e^{-t}\, t^{(n+\frac{1}{2})-1}\, dt$$

$$\Rightarrow \qquad \mu_{2n} = \frac{2^n \sigma^{2n}}{\sqrt{\pi}} \cdot \Gamma(n+\tfrac{1}{2})$$

Changing $n$ to $(n-1)$, we get

$$\mu_{2n-2} = \frac{2^{n-1} \cdot \sigma^{2n-2}}{\sqrt{\pi}}\ \Gamma(n-\tfrac{1}{2})$$

$$\therefore \qquad \frac{\mu_{2n}}{\mu_{2n-2}} = 2\sigma^2 \cdot \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n-\frac{1}{2})} = 2\sigma^2 (n-\tfrac{1}{2}) \ [\because \Gamma(r) = (r-1)\,\Gamma(r-1)]$$

$$\Rightarrow \qquad \mu_{2n} = \sigma^2 (2n-1)\, \mu_{2n-2} \qquad \ldots(8\cdot11)$$

which gives the *recurrence relation* for the moments of normal distribution.

From (8·11), we have

$$\mu_{2n} = [(2n-1)\sigma^2]\,[(2n-3)\sigma^2]\,\mu_{2n-4}$$

$$= [(2n-1)\sigma^2]\,[(2n-3)\sigma^2]\,[(2n-5)\sigma^2]\,\mu_{2n-6}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$= [(2n-1)\sigma^2]\,[(2n-3)\sigma^2]\,[(2n-5)\sigma^2] \ldots (3\sigma^2)(1\sigma^2) \cdot \mu_0$$

$$= 1.3.5 \ldots (2n-1)\,\sigma^{2n} \qquad \ldots(8\cdot12)$$

From (8·10) and (8·12) we conclude that *for the normal distribution all odd order moments about mean vanish and the even order moments about mean are given by (8·12).*

**Aliter.** The above result can also be obtained quite conveniently as follows:

The m.g.f. (about mean) is given by

$$E[e^{t(X-\mu)}] = e^{-\mu t} E(e^{tX}) = e^{-\mu t} M_X(t)$$

where $M_X(t)$ is the m.g.f. (about origin).

$$\therefore \text{ m.g.f. (about mean)} = e^{-\mu t} e^{\mu t + t^2\sigma^2/2} = e^{t^2\sigma^2/2}$$

$$= \left[1 + (t^2\sigma^2/2) + \frac{(t^2\sigma^2/2)^2}{2!} + \frac{(t^2\sigma^2/2)^3}{3!} + \ldots + \frac{(t^2\sigma^2/2)^n}{n!} + \ldots\right] \quad \ldots(8\cdot13)$$

The co efficient of $\frac{t^r}{r!}$ in (8·13) gives $\mu_r$, the $r$th moment about mean. Since there is no term with odd powers of $t$ in (8·13), all moments of odd order about mean vanish.

*i.e.,* $\mu_{2n+1} = 0 ; n = 0, 1, 2, \ldots$

and $\mu_{2n} =$ Coefficient of $\frac{t^{2n}}{(2n)!}$ in (8·13) $= \frac{\sigma^{2n} \times (2n)!}{2^n n!}$

$$= \frac{\sigma^{2n}}{2^n n!} \cdot \left[2n(2n-1)(2n-2)(2n-3)\ldots 5.4.3.2.1\right]$$

$$= \frac{\sigma^{2n}}{2^n . n!}\left[1.3.5\ldots(2n-1)\right]\left[2.4.6\ \ldots\ (2n-2).2n\right]$$

$$= \frac{\sigma^{2n}}{2^n . n!}\left[1.3.5\ \ldots\ (2n-1)\right]2^n\left[1.2.3\ldots n\right]$$

$$= 1.3.5\ldots(2n-1)\,\sigma^{2n}$$

**Remark.** In particular, we have from (8·10) and (8·12),

$$\mu_3 = 0 \text{ and } \mu_2 = 1\,.\,\sigma^2, \mu_4 = 1\cdot3\ \sigma^4$$

Hence $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0$ and $\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\sigma^4}{\sigma^4} = 3,$

the results which have already been obtained in (8·9).

**8·2·8. A linear combination of independent normal variates is also a normal variate.** Let $X_i$, $(i = 1, 2, \ldots, n)$ be $n$ independent normal variates with mean $\mu_i$ and variance $\sigma_i^2$ respectively. Then

$$M_{X_i}(t) = \exp\left\{\mu_i t + (t^2\sigma_i^2/2)\right\} \quad \ldots(8\cdot14)$$

The m.g.f. of their linear combination $\sum_{i=1}^{n} a_i X_i$, where $a_1, a_2, \ldots, a_n$ are constants, is given by

$$M_{\Sigma a_i X_i}(t) = M_{a_1 X_1 + a_2 X_2 + \ldots + a_n X_n}(t)$$

$$= M_{a_1 X_1}(t) \, . \, M_{a_2 X_2}(t) \ldots M_{a_n X_n}(t)$$

( $\because$ $X_i$'s are independent )

$$= M_{X_1}(a_1 t) \, . \, M_{X_2}(a_2 t) \ldots M_{X_n}(a_n t) \qquad \ldots(8 \cdot 15)$$

$[\because M_{cX}(t) = M_X(ct)]$

From (8·14), we have

$$M_{X_i}(a_i t) = e^{\mu_i a_i t + t^2 a_i^2 \sigma_i^2 / 2}$$

$\therefore$ (8·15), gives

$$M_{\sum_i a_i X_i}(t) = [e^{\mu_1 a_1 t + t^2 a_1^2 \sigma_1^2/2} \times e^{\mu_2 a_2 t + t^2 a_2^2 \sigma_2^2/2} \times \ldots \times e^{\mu_n a_n t + t^2 a_n^2 \sigma_n^2/2}]$$

$$= \exp\left[\left(\sum_{i=1}^{n} a_i \mu_i\right) t + t^2 \left(\sum_{i=1}^{n} a_i^2 \sigma_i^2\right) \Big/ 2\right],$$

which is the m.g.f. of a normal variate with mean $\sum_{i=1}^{n} a_i \mu_i$ and variance $\sum_{i=1}^{n} a_i^2 \cdot \sigma_i^2$. Hence by uniqueness theorem of m.g.f.,

$$\sum_{i=1}^{n} a_i X_i \sim N\left[\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right]. \qquad \ldots(8 \cdot 15\, a)$$

**Remarks 1.** If we take $a_1 = a_2 = 1$, $a_3 = a_4 = \ldots = 0$, then

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

If we take $a_1 = 1$, $a_2 = -1$, $a_3 = a_4 = \ldots = 0$, then

$$X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

Thus we see that the sum as well as the difference of two independent normal variates is also a normal variate. This result provides a sharp contrast to the Poisson distribution, in which case though the sum of two independent Poisson variates is a Poisson variate, the difference is not a Poisson variate.

**2.** If we take

$a_1 = a_2 = \ldots = a_n = 1$, then we get $\qquad \ldots(8 \cdot 15\, b)$

$$\sum_{i=1}^{n} X_i \sim N\left[\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right]$$

*i.e.*, the sum of independent normal variates is also a normal variate, which establishes the *additive property* of the normal distribution.

**3.** If $X_i$; $i = 1, 2, \ldots, n$ are identically and independently distributed as $N(\mu, \sigma^2)$ and if we take $a_1 = a_2 = \ldots = a_n = 1/n$,

then $$\frac{1}{n} \sum_{i=1}^{n} X_i \sim N\left\{\frac{1}{n} \sum_{i=1}^{n} \mu, \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2\right\}$$

$$\Rightarrow \qquad \bar{X} \sim N(\mu, \sigma^2/n), \text{ where } \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

This leads to the following important conclusion :

If $X_i$, $(i = 1, 2, ..., n)$, are identically and independently distributed normal variates with mean $\mu$ and variance $\sigma^2$, then their mean $\bar{X}$ is also $N(\mu, \sigma^2/n)$.

**8·2·9. Points of Inflexion of Normal Curve.** At the point of inflexion of the normal curve, we should have

$$f''(x) = 0, \text{ and } f'''(x) \neq 0$$

For normal curve, we have from (8·6)

$$f''(x) = -\frac{f(x)}{\sigma^2}\left[1 - \frac{(x-\mu)^2}{\sigma^2}\right],$$

$$\therefore \qquad f''(x) = 0 \quad \Rightarrow \quad 1 - \frac{(x-\mu)^2}{\sigma^2} = 0 \quad \Rightarrow \quad x = \mu \pm \sigma$$

It can be easily verified that at the points $x = \mu \pm \sigma$, $f'''(x) \neq 0$.

Hence the points of inflexion of the normal curve are given by $x = \mu \pm \sigma$ and $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-1/2}$ *i.e.*, they are equi-distant (at a distance $\sigma$) from the mean.

**8·2·10. Mean Deviation from the Mean for Normal Distribution.**

$$\text{M.D. (about mean)} = \int_{-\infty}^{\infty} |x - \mu| f(x)\, dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} |x - \mu| \, e^{-(x-\mu)^2/2\sigma^2}\, dx$$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z| \, e^{-z^2/2}\, dz \qquad \left[\frac{x-\mu}{\sigma} = z\right]$$

$$= \frac{2\sigma}{\sqrt{2\pi}} \cdot \int_{0}^{\infty} |z| \, e^{-z^2/2}\, dz,$$

since the integrand $|z| e^{-z^2/2}$ is an even function of $z$.

Since in $[0, \infty]$, $|z| = z$, we have

$$\text{M.D. (about mean)} = \sqrt{2/\pi}\;\sigma \int_{0}^{\infty} z\, e^{-z^2/2}\, dz$$

$$= \sqrt{2/\pi}\;\sigma \int_{0}^{\infty} e^{-t}\, dt \qquad \left[\frac{z^2}{2} = t\right]$$

$$= \sqrt{2/\pi}\ \sigma \left| \frac{e^{-t}}{-1} \right|_0^{\infty}$$

$$= \sqrt{2/\pi}\ \sigma$$

$$= \frac{4}{5}\sigma \text{ (approx.)}$$

**8·2·11. Area Property (Normal Probability Integral).** If $X \sim N(\mu, \sigma^2)$, then the probability that random value of $X$ will lie between $X = \mu$ and $X = x_1$ is given by

$$P(\mu < X < x_1) = \int_{\mu}^{x_1} f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{x_1} e^{-(x-\mu)^2/(2\sigma^2)}\,dx$$

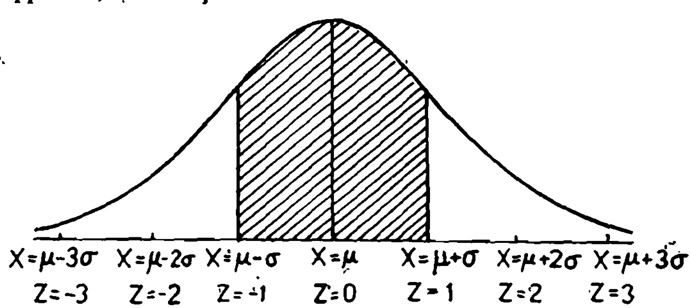
Put $\dfrac{X-\mu}{\sigma} = Z$, *i.e.*, $X - \mu = \sigma Z$

When $X = \mu$, $Z = 0$ and when $X = x_1$, $Z = \dfrac{x_1 - \mu}{\sigma} = z_1$, (say).

$$\therefore\quad P(\mu < X < x_1) = P(0 < Z < z_1) = \frac{1}{\sqrt{2\pi}} \int_0^{z_1} e^{-z^2/2}\,dz = \int_0^{z_1} \varphi(z)\,dz$$

where $\varphi(z) = \dfrac{1}{\sqrt{2\pi}} e^{-z^2/2}$, is the probability function of standard normal variate.

The definite integral $\int_0^{z_1} \varphi(z)\,dz$ is known as *normal probability integral* and gives the area under standard normal curve between the ordinates at $Z = 0$ and $Z = z_1$. These areas have been tabulated for different values of $z_1$, at intervals of 0·01 [c.f. Appendix, Table IV].




In particular, the probability that a random value of $X$ lies in the interval $(\mu - \sigma, \mu + \sigma)$ is given by

$$P(\mu - \sigma < X < \mu + \sigma) = \int_{\mu-\sigma}^{\mu+\sigma} f(x)\,dx$$

$$\Rightarrow \qquad P(-1 < Z < 1) = \int_{-1}^{1} \varphi(z)\, dz \qquad \left[ z = \frac{x - \mu}{\sigma} \right]$$

$$= 2\int_{0}^{1} \varphi(z)\, dz \qquad \text{(By symmetry)}$$

$$= 2 \times 0{\cdot}3413 = 0{\cdot}6826 \qquad \text{(From tables)} \quad \ldots(8{\cdot}17)$$

Similarly

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = P(-2 < Z < 2) = \int_{-2}^{2} \varphi(z)\, dz$$

$$= 2\int_{0}^{2} \varphi(z)\, dz = 2 \times 0{\cdot}4772 = 0{\cdot}9544 \quad \ldots(8{\cdot}18)$$

and

$$P(\sigma - 3\sigma < X < \mu + 3\sigma) = P(-3 < Z < 3) = \int_{-3}^{3} \varphi(z)\, dz$$

$$= 2\int_{0}^{3} \varphi(z)\, dz = 2 \times 0{\cdot}49865 = 0{\cdot}9973 \quad \ldots(8{\cdot}19)$$

Thus the probability that a normal variate $X$ lies outside the range $\mu \pm 3\sigma$ is given by

$$P(|X - \mu| > 3\sigma) = P(|Z| > 3) = 1 - P(-3 \leq Z \leq 3) = 0{\cdot}0027$$

Thus in all probability, we should expect a normal variate to lie within the range $\mu \pm 3\sigma$, though theoretically, it may range from $-\infty$ to $\infty$.

**Remarks. 1.** The total area under normal probability curve is unity, *i.e.*,

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{-\infty}^{\infty} \varphi(z)\, dz = 1$$

**2.** Since in the normal probability tables, we are given the areas under standard normal curve, in numerical problems we shall deal with the standard normal variate $Z$ rather than the variable $X$ itself.

**3.** If we want to find area under normal curve, we will somehow or other try to convert the given area to the form $P(0 < Z < z_1)$, since the areas have been given in this form in the tables.

**8·2·12. Error Function.** If $X \sim N(0, \sigma^2)$, then

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-x^2/2\sigma^2}, \quad -\infty < x < \infty$$

If we take $h^2 = \dfrac{1}{2\sigma^2}$ then $f(x) = \dfrac{h}{\sqrt{\pi}}\, e^{-h^2 x^2}$

The probability that a random value of the variate lies in the range $\pm x$ is given by

$$P = \int_{-x}^{x} f(x)\,dx = \frac{h}{\sqrt{\pi}} \int_{-x}^{x} e^{-h^2 x^2}\,dx$$

$$= \frac{2h}{\sqrt{\pi}} \int_{0}^{x} e^{-h^2 x^2}\,dx = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-h^2 x^2}\,(h dx) \qquad ...(*)$$

Taking

$\psi(y) = \frac{2}{\sqrt{\pi}} \int_{0}^{y} e^{-y^2}\,dy$, (*) may be re-written as

$$P = \psi(hx) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-h^2 x^2}\,(h dx) \qquad ...(**)$$

The function $\psi(y)$, known as the *error function*, is of fundamental importance in the *theory of errors* in Astronomy.

**8·2·3. Importance of Normal Distribution.** Normal distribution plays a very important role in statistical theory because of the following reasons :

(*i*) Most of the distributions occurring in practice, *e.g.*, Binomial, Poisson, Hypergeometric distributions, etc., can be approximated by normal distribution. Moreover, many of the sampling distributions, *e.g.*, Student's '*t*', Snedecor's F, Chi-square distributions, etc., tend to normality for large samples.

(*ii*) Even if a variable is not normally distributed, it can sometimes be brought to normal form by simple transformation of variable. For example, if the distribution of $X$ is skewed, the distribution of $\sqrt{X}$ might come out to be normal [c.f. Variate Transformations at the end of this Chapter].

(*iii*) If $X \sim N(\mu, \sigma^2)$, then

$$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0{\cdot}9973$$
$$\Rightarrow \quad P(-3 < Z < 3) = 0{\cdot}9973$$
$$\Rightarrow \quad P(|Z| < 3) = 0{\cdot}9973$$
$$\Rightarrow \quad P(|Z| > 3) = 0{\cdot}0027$$

This property of the normal distribution forms the basis of entire *Large Sample* theory.

(*iv*) Many of the distributions of sample statistic (*e.g.*, the distributions of sample mean, sample variance, etc.) tend to normality for large samples and as such they can best be studied with the help of the normal curves.

(*v*) The entire theory of small sample tests, *viz.*, $t$, $F$, $\chi^2$ tests-etc., is based on the fundamental assumption that the parent populations from which the samples have been drawn follow normal distribution.

(*vi*) Theory of normal curves can be applied to the graduation of the curves which are not normal.

(*vii*) Normal distribution finds large applications in Statistical Quality Control in industry for setting control limits.

The following quotation due to Lipman rightly reveals the popularity and importance of normal distribution.

*"Every body believes in the law of errors (the normal curve), the experimenters because they think it is a mathematical theorem, the mathematicians because they think it is experimental fact."*

W J Youden of the National Bureau of Standards describes the importance of the Normal distribution artistically in the following words :

THE NORMAL
LAW OF ERRORS
STANDS OUT IN THE
EXPERIENCE OF MANKIND
AS ONE OF THE BROADEST
GENERALISATIONS OF NATURAL
PHILOSOPHY IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES,
IN THE PHYSICAL AND SOCIAL SCIENCES
AND IN MEDICINE, AGRICULTURE AND
ENGINEERING. IT IS AN INDISPENSABLE TOOL FOR
THE ANALYSIS AND THE INTERPRETATION OF THE
BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT.

The above presentation, strikingly enough, gives the shape of the normal probability curve.

**8·2 14. Fitting of Normal Distribution.** In order to fit normal distribution to the given data we first calculate the mean $\mu$, (say), and standard deviation $\sigma$, (say), from the given data. Then the normal curve fitted to the given data is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-(x-\mu)^2/2\sigma^2\}, \quad -\infty < x < \infty$$

To calculate the expected normal frequencies we first find the standard normal variates corresponding to the lower limits of each of the class intervals, *i.e.*, we compute $z_i = \dfrac{x_i' - \mu}{\sigma}$, where $x_i'$ is the lower limit of the $i$th class interval. Then the areas under the normal curve to the left of the ordinate at $z = z_i$, say, $\varphi(z_i)$ are computed from the tables. Finally, the areas for the successive class intervals are obtained by subtraction, *viz.*, $\varphi(z_{i+1}) - \varphi(z_i)$, $(i = 1, 2, \ldots)$ and on multiplying these areas by $N$, we get the expected normal frequencies.

**Example 8·10.** *Obtain the equation of the normal curve that may be fitted to the following data :*

| *Class.* | 60–65 | 65–70 | 70–75 | 75–80 | 80–85 | 85–90 | 90–95 | 95–100 |
|---|---|---|---|---|---|---|---|---|
| *Frequency.* | 3 | 21 | 150 | 335 | 326 | 135 | 26 | 4 |

*Also obtain the expected normal frequencies*

**Solution.** For the given data, we have

$N = 1000$, $\mu = 79{\cdot}945$ and $\sigma = 5{\cdot}545$

Hence the equation of the normal curve fitted to the given data is

$$f(x) = \frac{1000}{\sqrt{2\pi} \times 5{\cdot}545} \exp\left\{ -\frac{1}{2}\left( \frac{x - 79{\cdot}945}{5{\cdot}545} \right)^2 \right\}$$

Theoretical normal frequencies can be obtained as follows .

| class | Lower class boundry ($X'$) | $Z = \frac{X' - \mu}{\sigma}$ | $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-z^2/2} dz$ | $\Delta\varphi(z) = \psi_{z+1} - \psi_z$ | Expected frequency $N\Delta\psi(z)$ |
|---|---|---|---|---|---|
| Below 60 | $-\infty$ | $-\infty$ | 0 | 0·000112 | 0·12 ≅ 0 |
| 60–65 | 60 | – 3·663 | 0·000112 | 0·002914 | 2·914 ≅ 3 |
| 65–70 | 65 | –2·745 | 0·003026 | 0·031044 | 31·044 ≅ 31 |
| 70–75 | 70 | –1·826 | 0·034070 | 0·147870 | 147·870 ≅ 148 |
| 75–80 | 75 | –0·908 | 0·181940 | 0·322050 | 322·050 ≅ 322 |
| 80–85 | 80 | 0·010 | 0·503990 | 0·319300 | 319·300 ≅ 319 |
| 85–90 | 85 | 0·928 | 0·823290 | 0·144072 | 144·072 ≅ 144 |
| 90–95 | 90 | 1·487 | 0·967362 | 0·029792 | 29·792 ≅ 30 |
| 95–100 | 95 | 2·675 | 0·997154 | 0·002733 | 2·733 ≅ 3 |
| 100 and over | 100 | 3·683 | 0·999887 | | |
| Total | | | | | 1000 |

**Example 8·11.** *For a certain normal distribution, the first moment about 10 is 40 and the fourth moment about 50 is 48. What is the arithmetic mean and standard deviation of the distribution ?*

**[Delhi Univ. B.Sc. (Hons. Subs.), 1987; Allahabad Univ. B.Sc. 1990]**

**Solution.** We know that if $\mu_1'$ is the first moment about the point $X = A$, then arithmetic mean is given by:

$$\text{Mean} = A + \mu_1'$$

We are given

$\mu_1'$ (about the point $X = 10$) $= 40 \Rightarrow$ Mean $= 10 + 40 = 50$

Also we are given

$\mu_4'$ (about the point $X = 50$) $= 48$, *i.e.*, $\mu_4 = 48$ $(\because \text{Mean} = 50)$

But for a normal distribution with standard deviation $\sigma$,

$$\mu_4 = 3\sigma^4 \Rightarrow 3\sigma^4 = 48 \quad i.e., \quad \sigma = 2$$

**Example 8·12.** *X is normally distributed and the mean of X is 12 and S.D. is 4. (a) Find out the probability of the following :*

(*i*) $X \geq 20$, (*ii*) $X \leq 20$, and (*iii*) $0 \leq X \leq 12$

(*b*) *Find* $x'$, *when* $P(X > x') = 0{\cdot}24$.

(*c*) *Find* $x_0'$ *and* $x_1'$, *when* $P(x_0' < X < x_1') = 0{\cdot}50$ *and* $P(X > x_1') = 0{\cdot}25$

**Solution.** (*a*) We have $\mu = 12$, $\sigma = 4$, *i.e.*, $X \sim N(12, 16)$.

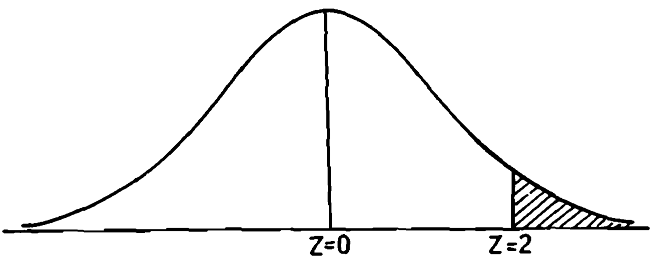
(i) $P(X \geq 20) = ?$

When $X = 20$, $Z = \dfrac{20-12}{4} = 2$

$\therefore \quad P(X \geq 20) = P(Z \geq 2) = 0{\cdot}5 - P(0 \leq Z \leq 2) = 0{\cdot}5 - 0{\cdot}4772 = 0{\cdot}0228$

(ii) $P(X \leq 20) = 1 - P(X \geq 20)$ ($\because$ Total probability = 1)

$= 1 - 0{\cdot}0228 = 0{\cdot}9772$



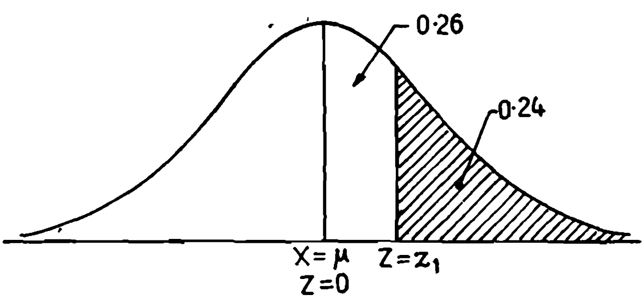

(iii) $P(0 \leq X \leq 12) = P(-3 \leq Z \leq 0)$ $\left(Z = \dfrac{X-12}{4}\right)$

$= P(0 \leq Z \leq 3) = 0{\cdot}49865$ (From symmetry)

(b) When $X = x'$, $Z = \dfrac{x'-12}{4} = z_1$ (say)

then, we are given

$P(X > x') = 0{\cdot}24 \quad \Rightarrow \quad P(Z > z_1) = 0{\cdot}24$, *i.e.*, $P(0 < Z < z_1) = 0{\cdot}26$




$\therefore$ From normal tables,

$z_1 = 0{\cdot}71$ (approx.)

Hence $\dfrac{x_1'-12}{4} = 0{\cdot}71 \quad \Rightarrow \quad x_1' = 12 + 4 \times 0{\cdot}71 = 14{\cdot}84$

(c) We are given

$P(x_0' < X < x_1') = 0{\cdot}50$ and $P(X > x_1') = 0{\cdot}25$ ...(*)