APPLIED STOCHASTIC PROCESSES

3 authors:

Christian Akrong Hesse
University of Professional Studies

John Benjamin Ofosu
Methodist University College Ghana

Frank Otchere
University of Ghana

APPLIED

STOCHASTIC PROCESSES

J. B. OFOSU, BSc, PhD, FSS

Professor of Statistics and Director, Quality Assurance Unit

Methodist University College Ghana

C. A. HESSE, BSc, MPhil

Lecturer in Statistics

Methodist University College Ghana

+233 244 648 757

FRANK OTCHERE, Bsc, MA, MPhil

Assistant Research Fellow

Institute of Statistical, Social and Economic Research,

University of Ghana

Ofosu, J. B., Hesse, C. A. & Otchere, F. (2014). Applied stochastic processes. *EPP Books Services, Accra*.

PREFACE

This book began many years ago, as lecture notes for students at King Saud University in Saudi Arabia, and later at the Methodist University College Ghana. Students in their third year of their undergraduate study, have used draft versions of the chapters, which have been revised and enriched.

The book is intended as a beginning text in stochastic processes for students familiar with elementary probability theory. The objectives of the book are threefold:

1. To introduce students to use standard concepts and methods of stochastic process.
2. To illustrate the diversity of applications of stochastic processes.
3. To provide exercises in the application of simple stochastic analysis to appropriate problems.

The book contains eight chapters. In Chapter 1, we review elementary probability theory. In particular, we cover the notion of conditional expectation, which is very useful in the sequel. The main characteristics of stochastic processes are given in Chapter 2. Important properties such as the concepts of independence and stationary increments are explained.

In Chapter 3, we discuss the class of Markov processes in discrete time with discrete state spaces and in Chapter 4 we consider asymptotic behaviour of Markov chains.

In Chapter 5, we discuss the classification of the states of Markov chains and in Chapter 6, we discuss birth-and-death processes.

In Chapter 7, the Poisson process, which is probably the most important stochastic process for students in telecommunications, is studied in detail. Several generalizations of this process, including non homogenous Poisson processes can be found in this Chapter.

Finally, Chapter 8 is concerned with the theory of queues. The models with a single server and those with at least two servers are treated separately. In general, we limit ourselves to the case of exponential models, in which both the times between arrivals of successive customers and the service times are exponential random variables. This chapter then becomes an application of Chapter 7.

The main pedagogical features of the book are threefold:

1. Each Chapter has an extensive collection of exercises, including end-of-section exercises that emphasize the material in the section. Most of the exercises are from published sources, including past examination questions from King Saud University and Methodist University College Ghana. Answers to all the exercises are given at the end of the book.
2. References cited in each Chapter are listed at the end of the Chapter.
3. Includes entertaining mini-biographies of mathematicians who have contributed to the theory, giving an enriching historical context.

We are indebted to a large number of people in the production of this book. We particularly appreciate students and professional practitioners who provided feedback, often in the form of penetrating questions that led to rewriting or expansion of the material in the book.

We thank King Saud University and Methodist University College Ghana, for permission to use their past examination questions in Statistics.

We have discussed some technical issues with a number of people, including Professors Lotfi Tadj, A. M. Aboummoh and A. Al-Zaid. To them we are most grateful.

Last, but not least, we wish to express our appreciation to Ms Wendy Linda Wordie of EPP Books Services, for her continued encouragement.

J. B. Ofosu

C. A. Hesse

F. Otchere

Accra

May, 2013

CONTENTS

FOREWORD

PROFESSOR J. ROBERTS,
UNIVERSITY OF CALIFORNIA, USA

The book is one of the most readable texts on stochastic process. I shall definitely recommend it to my students.

PROFESSOR G. WETHERITH,
UNIVERSITY OF COLORADO, USA

The writing style is unquestionably a strength of this book, particularly when compared to competing books.

PROFESSOR J. M. HILLER,
UNIVERSITY OF KUWAIT

This is a superb text from which to teach stochastic processes. The chapters are as far as possible self-contained and each contains material of varied difficulty, starting always with simple examples of stochastic process.

It is carefully written and illustrating account of stochastic processes, written at a level which will make it useful to a large class of readers.

NOTABLE FEATURES

Includes numerous real-world examples from physical and biological sciences.

Provides extensive collection of exercises for each chapter.

Includes useful references.

Answers to all the exercises are given at the end of the book.

Includes entertaining, mini-biographies of mathematicians, giving enriching historical context.

CHAPTER ONE

Review of Probability Theory

1.1 Elementary probability

Definition 1.1 (Random experiment)

A ***random experiment*** is an experiment that can be repeated under the same conditions and whose outcomes cannot be predicted with certainty.

Example 1.1

Picking a ball from a box containing 20 numbered balls, is a random experiment, since the 20 possible outcomes of the experiment cannot be predicted with certainty.

Example 1.2

Consider the experiment of rolling a six-sided die and observing the number which appears on the uppermost face of the die. The result can be any of the numbers 1, 2, 3, …, 6. This is a random experiment, since the outcome is uncertain.

Example 1.3

If we measure the distance between two points A and B, many times, under the same conditions, we expect to have the same result. This is therefore not a random experiment. It is a ***deterministic experiment***. If a deterministic experiment is repeated many times under exactly the same conditions, we expect to have the same result.

Definition 1.2 (*The sample space of an experiment*)

The ***sample space S*** of a random experiment is the set of all possible outcomes of the experiment.

Example 1.4

Define a sample space for each of the following experiments.

(a) The heights, in centimetres, of five children are 60, 65, 70, 45, 48. Select a child from this group of children, then measure and record the child's height.
(b) Select a number at random from the interval [0, 2] of real numbers. Record the value of the number selected.

Solution
(a) $S = \{60, 65, 70, 45, 48\}$. (b) $S = \{x: 0 \le x \le 2$, where x is a real number$\}$.

Definition 1.3 (Event)

An ***event*** is a subset of a sample space of an experiment. In particular, each possible outcome of a random experiment is called an ***elementary event***.

We often use [1]Venn diagrams in elementary probability: the sample space S is represented by a rectangle and the events A, B, C, etc., are represented by circles that overlap inside the rectangle (see Fig. 1.1).

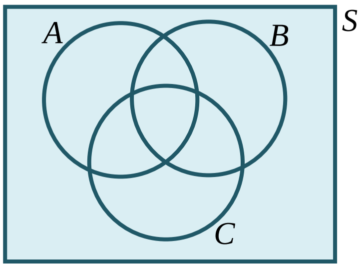

Fig. 1.1: *A Venn diagram*

Operations on events
1. $A \cup B$ denotes the event "A or B or both".
2. $A \cap B$ denotes the event "both A and B".
3. A' (or A^c) denotes the event which occurs if and only if A does not occur.

Mutually exclusive (or disjoint) events
Any two events A and B, are said to be mutually exclusive if $A \cap B = \phi$.

De Morgan's laws
The following theorems can be proved by means of Venn diagrams.

Theorem 1.1

$$(A \cup B)' = A' \cap B'.$$

Theorem 1.2

$$(A \cap B)' = A' \cup B'.$$

[1] John Venn, 1834 – 1923, was born and died in England. He was a mathematician and a priest. He taught at the University of Cambridge and worked in both mathematical logic and probability theory.

Theorem 1.3

If A, B and C are events defined on the same sample space, then:
$$(A \cup B \cup C)' = A' \cap B' \cap C', \quad (A \cap B \cap C)' = A' \cup B' \cup C', \quad (A' \cup B' \cup C')' = A \cap B \cap C.$$

The above results are called de Morgan's laws.

Counting sample points

Theorem 1.4 (*The multiplication theorem*)

If an operation can be performed in n_1 ways and after it is performed in any one of these ways, a second operation can be performed in n_2 ways and, after it is performed in any one of these ways, a third operation can be performed in n_3 ways, and so on for k operations, then the k operations can be performed together in $n_1 n_2 \ldots n_k$ ways.

Example 1.5

How many even three-digit numbers can be formed from the digits 3, 2, 5, 6, and 9 if each digit can be used only once?

Solution

Since the number must be even, we have $n_1 = 2$ choices for the units position. For each of these, we have $n_2 = 4$ choices for the tens position and then $n_3 = 3$ choices for the hundreds position. Therefore we can form a total of
$$n_1 \times n_2 \times n_3 = 2 \times 4 \times 3 = 24$$
even three-digit numbers.

Theorem 1.5 (*Permutations of n different things taken k at a time*)

If k objects are drawn at random from n distinct objects and if the order in which the objects are drawn does not matter, then the number of different permutations that can be obtained is
$$\underbrace{n \times n \times \ldots \times n}_{k \text{ times}} = n^k,$$
if the objects are taken ***with replacement*** and by
$$n \times (n-1) \times \ldots \times [n-(k-1)] = {}^nP_k = \frac{n!}{(n-k)!}, \qquad k = 0, 1, \ldots, n,$$
when the objects are taken ***without replacement***.

Example 1.6

Two lottery tickets are drawn from 20 for a first and a second prize. In how many ways can this be done?

Solution

The total number of ways is

$$^{20}P_2 = \frac{20!}{(20-2)!} = \frac{20!}{18!} = \frac{20 \times 19 \times 18!}{18!} = 20 \times 19 = 380.$$

Theorem 1.6 (*Permutations with repetitions*)

Given a set of n objects having n_1 elements alike of one kind, and n_2 elements alike of another kind, and n_3 elements alike of a third kind, and so on for k kinds of objects, the number of different arrangements of the n objects, taken all together is $\binom{n}{n_1, n_2, \ldots, n_k} = \dfrac{n!}{n_1!n_2!\ldots n_k!}$, where $n_1 + n_2 + \ldots + n_k = n$.

Example 1.7

How many different ways can 3 red, 4 yellow, and 2 blue bulbs be arranged in a string of Christmas tree lights with 9 sockets?

Solution

The total number of distinct arrangements is

$$\binom{9}{3,4,2} = \frac{9!}{3!\,4!\,2!} = 1\,260.$$

Combinations

In many problems, we are interested in the number of ways of selecting r objects from n without regard to order. These selections are called ***combinations***.

Theorem 1.7

The number of combinations of n distinct objects taken r at a time is

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}.$$

Example 1.8

From 4 chemists and 3 physicists, find the number of committees that can be formed consisting of 2 chemists and 1 physicist.

Solution

The number of ways of selecting 2 chemists from 4 is $\binom{4}{2} = \frac{4!}{2!\,2!} = 6.$

The number of ways of selecting 1 physicist from 3 is $\binom{3}{1} = \frac{3!}{1!\,2!} = 3.$

Using the multiplication theorem (see Theorem 1.4 on page 3) with $n_1 = 6$ and $n_2 = 3$, it can be seen that we can form $n_1 \times n_2 = 6 \times 3 = 18$ committees with 2 chemists and 1 physicist.

1.2 Some probability laws

Axioms of probability

Let S be the sample space of an experiment and P a set function which assigns a number $P(A)$ to every $A \subset S$. Then the function $P(A)$ is said to be a probability function if it satisfies the following three axioms:

***Axiom* 1**: $P(S) = 1.$

***Axiom* 2**: $P(A) \geq 0$ for every event A.

***Axiom* 3**: If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$.

Definition 1.4 (Classical definition of *P*(*A*))

If all the simple events in S are equally likely, then

$$P(A) = \frac{n(A)}{n(S)} \quad \text{for all } A \subset S,$$

where $n(A)$ denotes the number of elements in A.

This classical definition dates back to the 17^{th} century and the work of Pascal and Fermat [see Todhunter (1931) and David (1962)].

Elementary Theorems

The following theorems can be proved from the axioms of probability.

Theorem 1.8

$$P(\phi) = 0.$$

Theorem 1.9

$$P(A') = 1 - P(A).$$

Sometimes it is more difficult to calculate the probability that an event occurs than it is to calculate the probability that the event does not occur. Should this be the case for some event A, we first find $P(A')$ and then, using Theorem 1.9, we find $P(A)$ by subtraction.

Theorem 1.10

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Theorem 1.10 is often called the ***addition rule of probability***.

Corollary 1.1

If the events A and B are mutually exclusive, then $A \cap B = \phi$ and so by Theorem 1.8, $P(A \cap B) = 0$. Theorem 1.10 then becomes

$$P(A \cup B) = P(A) + P(B). \quad \text{...(1.2.1)}$$

Corollary 1.2

If the events $A_1, A_2, ..., A_n$ are mutually exclusive, then:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i). \quad \text{...(1.2.2)}$$

The following corollary gives an extension of Theorem 1.10 to 3 events.

Corollary 1.3

If A, B, and C are three events defined on the same sample space, then:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

Example 1.9

The following table shows 100 patients classified according to blood group and sex.

	Blood group		
	A	*B*	*O*
Male	30	20	17
Female	15	10	8

If a patient is selected at random from the 100 patients, find the probability that the patient selected:

(a) is a male or has blood group A, (b) does not have blood group A,

(c) is a female or does not have blood group B.

Solution

There are 100 ways in which we can select a patient from the 100 patients. Since the patient is selected at random, all the 100 ways of selecting a patient are equally likely.

(a) Let M denote the event "a patient selected is a male" and A the event "a patient selected has blood group A". We wish to find $P(M \cup A)$. By the addition rule of probability,

$$\begin{aligned} P(M \cup A) &= P(M) + P(A) - P(M \cap A) \\ &= \frac{n(M)}{100} + \frac{n(A)}{100} - \frac{n(M \cap A)}{100} \\ &= \frac{67}{100} + \frac{45}{100} - \frac{30}{100} = \frac{82}{100} = 0.82. \end{aligned}$$

(b) We wish to find $P(A')$. By Theorem 1.9,

$$P(A') = 1 - P(A) = 1 - \frac{n(A)}{100} = 1 - \frac{45}{100} = 0.55.$$

(c) Let F denote the event "a patient selected is a female" and B the event "a patient selected has blood group B". We wish to find $P(F \cup B')$. By the addition rule of probability,

$$\begin{aligned} P(F \cup B') &= P(F) + P(B') - P(F \cap B') \\ &= P(F) + \{1 - P(B)\} - P(F \cap B') \\ &= \frac{33}{100} + \left(1 - \frac{30}{100}\right) - \frac{23}{100} = 1 - \frac{20}{100} = 0.8. \end{aligned}$$

Example 1.10

The probability that a new airport will get an award for its design is 0.04; the probability that it will get an award for the efficient use of materials is 0.2 and the probability that it will get both awards is 0.03. Find the probability that it will get:

(a) at least one of the two awards, (b) only one of the two awards,

(c) none of the two awards.

Solution

Let D denote the event "the airport will get an award for its design", and E the event "the airport will get an award for the efficient use of materials".

We are given that $P(D) = 0.04$, $P(E) = 0.2$ and $P(D \cap E) = 0.03$. We can therefore draw Fig. 1.2. Notice that, since $P(D) = 0.04$, and $P(D \cap E) = 0.03$, $P(D \cap E') = 0.04 - 0.03 = 0.01$.

(a) We wish to find $P(D \cup E)$. From Fig. 1.2

(b) The probability that it will get only one of the awards is

$$P(D \cap E') + P(D' \cap E) = 0.01 + 0.17 = 0.18.$$

(c) We wish to find $P(D' \cap E')$. From Fig. 1.2,

$$P(D' \cap E') = 0.79.$$

Alternatively, using Theorems 1.2 and 1.9, we obtain

$$\begin{aligned} P(D' \cap E') &= 1 - P\left[(D' \cap E')'\right] \\ &= 1 - P(D \cup E) = 1 - 0.21, \quad \text{(from part (a))} \\ &= 0.79. \end{aligned}$$

1.3 Conditional probability

Notation

The expression $P(A|B)$ denotes the probability of the event A, given that (or knowing that, or simply if) the event B has occurred. We call it the ***conditional probability of A given B***. We now give a formal definition of conditional probability.

Definition 1.5

In particular, if S is a finite, equiprobable sample space, then:

$$P(A\cap B)=\frac{n(A\cap B)}{n(S)},\quad P(B)=\frac{n(B)}{n(S)},\text{ and so}$$

$$P(A|B)=\left\{\frac{n(A\cap B)}{n(S)}\right\}\Big/\left\{\frac{n(B)}{n(S)}\right\} = \frac{n(A\cap B)}{n(B)}. \quad \text{.................................(1.3.1)}$$

Example 1.11

Consider the data given in Example 1.9 on page 7. If a patient is chosen at random from the 100 patients, find the probability that the patient chosen has blood group A given that he is a male.

Solution

Let M denote the event "the patient chosen is a male" and A the event "the patient chosen has blood group A". We wish to find $P(A|M)$. Using Definition 1.5, we obtain

$$P(A|M) = \frac{P(A\cap M)}{P(M)} = \frac{30/100}{67/100} = \frac{30}{67}.$$

Alternatively, since the sample space is finite and equiprobable, by Equation (1.3.1),

$$P(A|M) = \frac{n(A\cap M)}{n(M)} = \frac{30}{67}.$$

1.4 The multiplication rule

If, in an experiment, the events A and B can both occur, then

$$P(A\cap B)=P(A)P(B|A)=P(B)P(A|B) \quad \text{...(1.4.1)}$$

This result is called the ***multiplication rule***. The multiplication rule can be applied to two or more events. For three events A, B and C, the multiplication rule takes the form

$$P(A\cap B\cap C) = P(A)P(B|A)P(C|A\cap B). \quad \text{.....................................(1.4.2)}$$

where $P(A)\neq 0$ and $P(A\cap B)\neq 0$. The multiplication rule can be extended by mathematical induction to the following theorem.

Theorem 1.11

For any events $A_1, A_2, ..., A_n$ $(n>2)$

$$P(A_1\cap A_2\cap...\cap A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1\cap A_2)...P(A_n|A_1\cap A_2...\cap A_{n-1}).$$

Example 1.12

A box contains 5 red, 4 white and 3 blue balls. If three balls are drawn successively from the box, find the probability that they are drawn in the order red, white and blue if each ball is not replaced.

Solution

Let R be the event "red on first draw", W the event "white on second draw" and B the event "blue on third draw". We wish to find $P(R \cap W \cap B)$. Since there are 5 balls out of 12 balls, $P(R) = \frac{5}{12}$. If the first ball drawn is red, then there are 4 white balls out of the 11 balls remaining in the box. Hence $P(W \mid R) = \frac{4}{11}$. If the first ball is red and the second ball is white, then there are 3 blue balls out of the 10 balls remaining in the box. It follows that $P(B \mid R \cap W) = \frac{3}{10}$. Hence by Theorem 1.11,

$$P(R \cap W \cap B) = \left(\frac{5}{12}\right)\left(\frac{4}{11}\right)\left(\frac{3}{10}\right) = \frac{1}{22}.$$

A partition of a sample space

Definition 1.6

The events $B_1, B_2, \ldots, B_n$ constitute a partition of the sample space S if

(i) $B_i \cap B_j = \phi, \quad \forall\, i \neq j$, (ii) $\bigcup_{k=1}^{n} B_k = S$, (iii) $P(B_k) > 0, \quad k = 1, 2, \ldots, n.$

If $B_1, B_2, \ldots, B_n$ is a partition of a sample space S, then we may write, for any event A,

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_n) \quad \text{.......................................(1.4.3)}$$

where $(A \cap B_i) \cup (A \cap B_j) = \phi, \quad \forall\, i \neq j$. Making use of Axiom 3 in the definition of the function P, we obtain the following result.

Theorem 1.12 (total probability rule)

If $A \subset S$ and the events $B_1, B_2, \ldots, B_n$ form a partition of S, then

$$P(A) = \sum_{k=1}^{n} P(A \cap B_k) = \sum_{k=1}^{n} P(A \mid B_k) P(B_k). \quad \text{...........................(1.4.4)}$$

This result is called the ***total probability rule***. We deduce from the total probability rule and from the formula

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \text{ if } P(B)>0, \quad \ldots\ldots(1.4.5)$$

the result known as [2]Bayes' rule (or formula, or theorem).

Theorem 1.13 (Bayes' rule)

If $P(A)>0$, then

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{k=1}^{n} P(A|B_k)P(B_k)}, \text{ for } j=1,\ 2,\ \ldots,\ n \quad \ldots\ldots(1.4.6)$$

where $B_1,\ B_2,\ \ldots,\ B_n$ is a partition of S.

Bayes' theorem is applicable in situations where quantities of the form $P(A|B_i)$ and $P(B_i)$ are known and we wish to find $P(B_i|A)$. The following example illustrates an application of the theorem.

Example 1.13

A consulting firm rents cars from three agencies: 30% from agency A, 20% from agency B and 50% from agency C. 15% of the cars from A, 10% of the cars from B and 6% of the cars from C have bad tyres. If a car rented by the firm has bad tyres, find the probability that it came from agency C.

Solution

Let A_1 denote the event "the car came from agency A",

A_2 denote the event "the car came from agency B",

A_3 denote the event "the car came from agency C",

and let T denote the event "a car rented by the firm has bad tyres". We wish to find $P(A_3|T)$. We are given $P(A_1)=0.3$, $P(A_2)=0.2$, $P(A_3)=0.5$, $P(T|A_1)=0.15$, $P(T|A_2)=0.1$ and

[2]The Reverend Thomas Bayes, 1702 – 1761, was born and died in England. His works on probability theory were published in a posthumous scientific paper in 1764.

$P(T|A_3) = 0.06$. A_1, A_2 and A_3 are mutually exclusive and $P(A_1)+P(A_2)+P(A_3)=1$, and so A_1, A_2 and A_3 form a partition of the sample space. Hence, by Bayes' theorem,

$$\begin{aligned} P(A_3|T) &= \frac{P(A_3)P(T|A_3)}{P(A_1)P(T|A_1)+P(A_2)P(T|A_2)+P(A_3)P(T|A_3)} \\ &= \frac{0.5\times 0.06}{0.3\times 0.15 \ + \ 0.2\times 0.1 \ + \ 0.5\times 0.06} \quad = \quad 0.3158. \end{aligned}$$

1.5 Independent events

Definition 1.7

Two events with nonzero probabilities are independent if and only if, any one of the following equivalent statements is true.

(a) $P(A|B) = P(A)$, (b) $P(B|A) = P(B)$, (c) $P(A\cap B) = P(A)P(B)$.

Two events that are not independent are said to be dependent. Usually, physical conditions under which an experiment is performed, will enable us to decide whether or not two or more events are independent. In particular, ***the outcomes of unrelated parts of an experiment can be treated as independent***.

Theorem 1.14

If A and B are independent, then so are A' and B, A and B' and A' and B'.

Example 1.14

A pair of fair dice is thrown twice. Find the probability of getting totals of 7 and 11.

Solution

Let A_1, A_2, B_1 and B_2 be the respective events that a total of 7 occurs on the first throw, a total of 7 occurs on the second throw, a total of 11 occurs on the first throw and a total of 11 occurs on the second throw. We are interested in the probability of the event $(A_1\cap B_2)\cup(A_2\cap B_1)$. It is clear that the events A_1, A_2, B_1 and B_2 are independent. Moreover, $A_1\cap B_2$ and $A_2\cap B_1$ are mutually exclusive events. Hence,

$$\begin{aligned} P[(A_1\cap B_2)\cup(A_2\cap B_1)] &= P(A_1\cap B_2) \ + \ P(A_2\cap B_1) \\ &= P(A_1)P(B_2) \ + \ P(A_2)P(B_1) \\ &= \left(\tfrac{6}{36}\right)\left(\tfrac{2}{36}\right) \ + \ \left(\tfrac{6}{36}\right)\left(\tfrac{2}{36}\right) \ = \ \tfrac{1}{54}. \end{aligned}$$

Occasionally, we must deal with more than two events. Again, the question arises: When are these events considered independent? The following definition answers this question by extending our previous definition to include more than two events.

Definition 1.8

Let $A_1, A_2, ..., A_n$ be a finite collection of events $(n \geq 3)$. These events are mutually independent if and only if the probability of the intersection of any 2, 3, …, n of these events is equal to the product of their respective probabilities.

Definition 1.9

The three events A_1, A_2, and A_3 are independent if and only if:

$$P(A_1 \cap A_2) = P(A_1)P(A_2), \qquad P(A_1 \cap A_3) = P(A_1)P(A_3),$$
$$P(A_2 \cap A_3) = P(A_2)P(A_3), \qquad P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3).$$

Example 1.15

A certain system is made of n components which function independently from one another. Let F_i denote the event that component i functions at time t_0, $i = 1, 2, ..., n$ and let F denote the event that the system functions at time t_0. Find $P(F)$ if the components are connected:

(a) in series, (b) in parallel.

Solution

Fig. 1.3(a) shows a series system while Fig. 1.3(b) shows a parallel system.

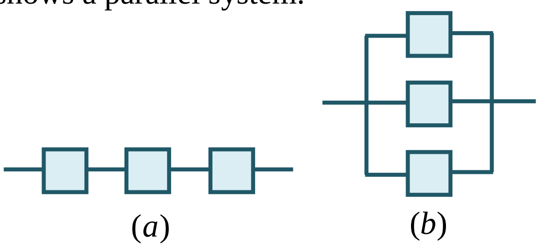

Fig. 1.3

(a) Here, $F = \bigcap_{i=1}^{n} F_i$ and so

$$P(F) = P\left(\bigcap_{i=1}^{n} F_i\right) = \prod_{i=1}^{n} P(F_i),$$

since $F_1, F_2, ..., F_n$ are independent.

(b) Here, $F = \bigcup_{i=1}^{n} F_i$ and so

$$P(F) = 1 - P(F') = 1 - P\left(\bigcap_{i=1}^{n} F_i'\right).$$

Since $F_1, F_2, ..., F_n$ are independent, $F_1', F_2', ..., F_n'$ are also independent.

Therefore, $P(F) = 1 - \left(\prod_{i=1}^{n} P(F_i')\right) = 1 - \prod_{i=1}^{n}\{1 - P(F_i)\}.$

Exercise 1(a)

1. If $A \subset B$, prove that $P(A) \le P(B)$. Hint: $B = A \cup (A' \cap B)$.

2. Prove that the probability of any event A is at most 1. Hint: $A \subseteq S$.

3. Let A and B be events with $P(A) = 0.25$, $P(B) = 0.40$ and $P(A \cap B) = 0.15$. Find
   (a) $P(A' \cap B')$, (b) $P(A \cap B')$, (c) $P(A' \cap B)$.

4. A certain carton of eggs has 3 bad and 9 good eggs.
   (a) If an omelette is made of 3 eggs randomly chosen from the carton, what is the probability that there are no bad eggs in the omelette?
   (b) What is the probability of having exactly 2 bad eggs in the omelette?

5. Samples of a cast aluminum part are classified on the basis of surface finish (in microinches) and length measurements. The results of 100 parts are summarized below.

		length	
		excellent	good
surface finish	excellent	75	7
	good	10	8

   Let A denote the event that a sample has excellent surface finish, and let B denote the event that a sample has excellent length. Find
   (a) $P(A)$, (b) $P(B)$, (c) $P(A')$, (d) $P(A \cap B)$, (e) $P(A \cup B)$.

6. Samples of foam from two suppliers are classified for conformance to specifications. The results from 40 samples are summarized below

		Conforms	
		yes	no
Supplier	1	18	2
	2	17	3

   Let A denote the event that a sample is from supplier 1, and let B denote the event that a sample conforms to specifications. Find

(a) $P(A)$, (b) $P(B)$, (c) $P(A')$, (d) $P(A \cap B)$, (e) $P(A \cup B)$, (f) $P(A' \cap B')$.

7. In a certain population of women, 4% have breast cancer, 20% are smokers and 3% are both smokers and have breast cancer. If a woman is selected at random from this population, find the probability that the person selected is:
   (a) a smoker or has breast cancer,
   (b) a smoker and does not have breast cancer,
   (c) not a smoker and does not have breast cancer.

8. A study of major flash floods that occurred over the last 15 years indicates that the probability that a flash flood warning will be issued is 0.5 and the probability of a dam failure during the flood is 0.33. The probability of a dam failure given that a warning is issued is 0.17. Find the probability that a flash flood warning will be issued and a dam failure will occur.

9. Show that if A_1 and A_2 are independent, then A_1 and A_2' are also independent. Hint: $A_1 = (A_1 \cap A_2) \cup (A_1 \cap A_2')$.

10. Kofi feels that the probability that he will get an A in the first Physics test is $\frac{1}{2}$ and the probability that he will get $A's$ in the first and second Physics tests is $\frac{1}{3}$. If Kofi is correct, what is the conditional probability that he will get an A in the second test, given that he gets an A in the first test?

11. In rolling 2 balanced dice, if the sum of the two values is 7, what is the probability that one of the values is 1?

12. A random sample of 200 adults are classified below by sex and their level of education attained.

Education	Male	Female
Elementary	38	45
High School	28	50
University	22	17

If a person is chosen at random from this group, find the probability that:
(a) the person is a male, given that the person has High School education,
(b) the person does not have a university degree, given that the person is a female.

13. In an experiment to study the relationship between hypertension and smoking habits, the following data were collected for 180 individuals.

	Non-smokers	Moderate smokers	Heavy smokers
Hypertension	21	36	30
No hypertension	48	26	19

If one of these individuals is selected at random, find the probability that the person is
(a) experiencing hypertension, given that he/she is a heavy smoker;
(b) a non-

(b) a person with a room having faulty plumbing was assigned accommodation at hotel *B*?

19. Suppose that at a certain accounting office, 30%, 25% and 45% of the statements are prepared by Mr. George, Mr. Charles and Mrs. Joyce, respectively. These employees are very reliable. Nevertheless, they are in error some of the time. Suppose that 0.01%, 0.005% and 0.003% of the statements prepared by Mr. George, Mr. Charles and Mrs. Joyce, respectively, are in error. If a statement from the accounting office is in error, what is the probability that it was prepared (caused) by Mr. George?

20. A certain construction company buys 20%, 30%, and 50% of their nails from hardware suppliers *A*, *B*, and *C*, respectively. Suppose it is known that 0.05%, 0.02% and 0.01% of the nails from *A*, *B*, and *C*, respectively, are defective.
    (a) What percentage of the nails purchased by the construction company are defective?
    (b) If a nail purchased by the construction company is defective, what is the probability that it came from the supplier *C*?

1.6 Random variables

1.6.1 Introduction

Definition 1.10 (*Random variable*)

A random variable is a function X that assigns a real number $X(s) = x$ to each element s of S, where S is the sample space of a random experiment (see Fig. 1.4). We denote by S_X the set of all values of X.

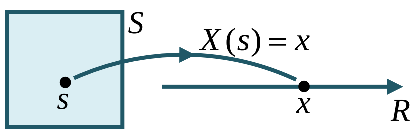

***Fig.* 1.4**

Example 1.16

Three balls are drawn in succession, without replacement, from a box containing 5 white and 4 green balls. Let Y denote the number of white balls selected. The possible outcomes of the experiment and the values y of Y, are:

outcome	*GGG*	*GGW*	*GWG*	*WGG*	*GWW*	*WGW*	*WWG*	*WWW*
y	0	1	1	1	2	2	2	3

where G denotes "green" and W denotes "white" and the i^{th} letter in a triple, denotes the colour of the i^{th} ball drawn (i = 1, 2, 3). For example, *GWG* means the first ball drawn is green, the second ball drawn is white and the third ball drawn is green.

Example 1.17

Let W denote the number of times a die is thrown until a 3 occurs. The possible outcomes and the values w, of the random variable W are:

outcome	F,	NF,	NNF,	$NNNF$, …
w	1,	2,	3,	4,…

where F and N represent, respectively, the occurrence, and non-occurrence of a 3. The random variable W takes the values 1, 2, 3, 4, … . The range of W is said to be ***countably infinite***.

Example 1.18

Let H denote the height, in metres, of a patient selected from a hospital. Values of H depend on the outcomes of the experiment. H is therefore a random variable.

In Example 1.16, the range of Y is finite while in Example 1.17, the range of W is countably infinite. The random variables Y and W are examples of discrete random variables.

It is easy to distinguish a discrete random variable from one that is not discrete. Just ask the question: "What are the possible values for the random variable?" If the answer is a finite set or a countably infinite set, then the random variable is discrete; otherwise, it is not. This idea leads to the following definition.

Definition 1.11 (Discrete random variable)

> A random variable is discrete if it can assume a finite or a countably infinite set of values.

In Example 1.18, the range of the random variable H is neither finite nor countably infinite. H can assume any value in some interval of real numbers. H is an example of a continuous random variable. We therefore have the following definition.

Definition 1.12 (Continuous random variable)

> If the range of a random variable X contains an interval (either finite or infinite) of real numbers, then X is a ***continuous random variable***.

In most practical problems, continuous random variables represent measured data, such as heights, temperatures and distances, whereas discrete random variables represent count data, such as number of road traffic accidents in Accra in a week.

Definition 1.13 (The distribution function)

The distribution function of a random variable X is defined by

$$F_X(x) = P(X \le x) \quad \forall\, x \in R.$$

Properties

(i) $0 \le F_X(x) \le 1$, (ii) $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to\infty} F_X(x) = 1$,

(iii) If $x_1 < x_2$, then $F_X(x_1) \le F_X(x_2)$ (iv) $P(a < X \le b) = F_X(b) - F_X(a)$.

1.6.2 Discrete random variables

Definition 1.14 (*Probability mass function*)

A function $p_X(x)$ is the probability mass function of a discrete random variable X if it has the following three properties:

(1) $p_X(x_i) = P(X = x_i)$, $\forall\, x_i \in S_X$, (2) $p_X(x_i) \ge 0$, $\forall\, x_i \in S_X$, (3) $\sum_{x_i \in S_X} p_X(x_i) = 1$.

It can be seen that, if X is a discrete random variable, then

$$F_X(x) = \sum_{x_i \le x} p_X(x_i). \qquad \text{(1.6.1)}$$

Exercise 1(b)

1. A discrete random variable X has a probability mass function given by
   $p_X(x) = c(x+1)$, $x = 0, 1, 2, 3$.
   (a) Find the value of the constant c. (b) Find: (i) $P(0 \le X < 2)$, (ii) $P(X > 1)$.
2. Determine whether each of the following functions can serve as a probability mass function of a discrete random variable X:
   (a) $p_X(x) = \frac{1}{2}(x-1)$, $x = 0, 1, 2, 3$.
   (b) $p_X(x) = \frac{1}{10}x$, $x = 1, 2, 3, 4$.
   (c) $p_X(x) = \frac{1}{6}x^2$, $x = -1, 0, 1, 2$.
3. Let X be a random variable whose probability mass function is defined by the values
   $p_X(-2) = \frac{1}{10}$, $p_X(0) = \frac{2}{10}$, $p_X(4) = \frac{4}{10}$, $p_X(11) = \frac{3}{10}$.

Find: (a) $P(-2 \le X < 4)$, (b) $P(X > 0)$, (c) $P(X \le 4)$.

4. Check whether the following functions satisfy the conditions of a probability mass function.

   (a) $p_X(x) = \frac{1}{4}$, $x = -3, 0, 1, 4$.

   (b) $p_X(x) = \frac{1}{x}$, $x = 1, 2, 3, 4$.

   (c) $p_X(x) = 1 - x$, $x = 0, \frac{1}{2}, \frac{3}{2}$.

   (d) $p_X(x) = \left(\frac{1}{2}\right)^x$, $x = 1, 2, 3, 4, \ldots$.

5. Consider a throw of two fair dice. Let X denote the sum of the numbers which appear on the uppermost faces of the two dice.

   (a) Find the probability mass function of X.

   (b) Find: (i) $P(X = 7)$, (ii) $P(X > 8)$, (iii) $P(3 < X < 11)$.

6. Determine the value of c so that each of the following can serve as a probability mass function of a discrete random variable.

   (a) $p_X(x) = c(x^2 + 4)$, $x = 0, 1, 2, 3$.

   (b) $p_X(x) = c\binom{2}{x}\binom{3}{3-x}$, $x = 0, 1, 2$.

7. A discrete random variable X has the probability mass function given by

$$p_X(x) = \begin{cases} a\left(\frac{1}{3}\right)^{x-1}, & x = 1, 2, \ldots \\ 0, & \text{elsewhere.} \end{cases}$$

   (a) Find the value of a. (b) Find $P(X = 3)$.

1.6.3 Important discrete random variables

(i) *The discrete uniform distribution*

Definition 1.15

A random variable X has the discrete uniform distribution with parameter n if it assumes the values $x_1, x_2, \ldots, x_n$ with equal probabilities. That is, if

$$p_X(x_i) = P\left(X = x_i\right) = \frac{1}{n}, \quad i = 1, 2, \ldots, n. \quad \ldots\ldots(1.6.2)$$

(ii) [3]*Bernoulli distribution*

Definition 1.16

A discrete random variable X is said to have the Bernoulli distribution with parameter p, where p is called the ***probability of success***, if its probability mass function is given by

$$p_X(x) = p^x(1-p)^{1-x}, \qquad x = 0, 1. \quad \text{.......................................}(1.6.3)$$

(iii) *The binomial distribution*

Definition 1.17

A random variable X has the binomial distribution with parameters n and p if

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, \qquad x = 0, 1, \ldots, n, \text{ where } q = 1 - p.$$

The probability mass function of the binomial distribution with parameters n and p is denoted by $b(x; n, p)$. The calculation of this probability mass function can be tedious when n is large. Fortunately, probabilities for different values of n and p have been tabulated. Table A.1, in the appendix, is one of many such tables.

Theorem 1.15

Let $Y_1, Y_2, \ldots, Y_n$ be independent Bernoulli random variables, each with parameter p. If $X = Y_1 + Y_2 + \ldots + Y_n$, then X has the binomial distribution with parameters n and p.

(iv) *The geometric (or [4]Pascal) distribution*

Definition 1.18

In a series of independent Bernoulli trials, with constant probability p of success, let the random variable X denote the number of trials until the first success. Then, X has the geometric distribution with parameter p, and

$$P(X = x) = p(1-p)^{x-1}, \qquad x = 1, 2, 3, \ldots$$

---

[3]Jacob (or Jacques) Bernoulli, 1654 – 1705, was born and died in Switzerland. His important book on probability theory was published eight years after his death.

[4] Blaise Pascal, 1623 – 1662, was born and died in France. He is one of the founders of the theory of probability. He was also interested in geometry and in physics, in addition to publishing books on philosophy and on theology.

(v) *The* [5]*Poisson distribution*

Definition 1.19

A random variable X has the Poisson distribution with parameter λ ($\lambda > 0$) if its probability mass function is given by

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad x = 0, 1, 2, 3, \ldots$$

Since the Poisson distribution has many applications, it has been tabulated. Table A.2, in the Appendix, gives the cumulative probabilities for selected values of λ.

Example 1.19

In the study of a certain aquatic organism, a large number of samples were taken from a pond, and the number of organisms in each sample was counted. The average number of organisms per sample was found to be two. Assuming that the number of organisms follows a Poisson distribution, find the probability that the next sample taken will contain

(a) one or fewer organisms, (b) exactly 3 organisms, (c) more than 5 organisms.

Solution

Let X denote the number of organisms in the next sample. Then X has the Poisson distribution with parameter $\lambda = 2$.

(a) We wish to find $P(X \le 1)$. From Table A.2, we see that when $\lambda = 2$,

$$P(X \le 1) = 0.406.$$

(b) We wish to find $P(X = 3)$. Now,

$$P(X = 3) = P(X \le 3) - P(X \le 2) = 0.857 - 0.677 = 0.180.$$

(c) We wish to find $P(X > 5)$. Now,

$$P(X > 5) = 1 - P(X \le 5) = 1 - 0.983 = 0.017.$$

[5]Siméon Denis Poisson, 1781 – 1840, was born and died in France. He first studied medicine and, from 1798, mathematics at the École Polytechnique de Paris, where he taught from 1802 to 1808. His professors at the École Polytechnique were, among others, Laplace and Lagrange. In mathematics, his main results were his papers on definite integrals and Fourier series. The Poisson distribution appeared in his important book on probability theory published in 1837. He also published works on mechanics, electricity, magnetism, and astronomy. His name is associated with numerous results in both mathematics and physics.

Theorem 1.16 (***The distribution of the sum of independent Poisson random variables***)

Let $X_1, X_2, \ldots, X_n$ be independent Poisson random variables with means $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. Then $Y = X_1 + X_2 + \ldots + X_n$ has the Poisson distribution with mean $\sum_{i=1}^{n} \lambda_i$.

Theorem 1.17 (***The Poisson approximation to the binomial distribution***)

Let X be a >] TJETBT29.24 15.1 5ad5q80.±28.w7B>E/F10 12 Tf1 0 0 1 103.7 620.02 Tm0

Example 1.20

Suppose that, on average, 1 person in 1 000 makes a numerical error in preparing his or her income tax return. If 8 000 forms are selected at random and examined, find the probability that less than 7 of the forms contain an error.

Solution

If X of the 8 000 forms contain an error, then X has the binomial distribution with parameters n = 8 000 and p = 0.001. Hence,

$$P(X < 7) = \sum_{x=0}^{6} \binom{8000}{x} (0.001)^x (0.999)^{8000-x}.$$

Binomial distribution tables do not exist for calculating this probability. Moreover, it is cumbersome to calculate the probabilities. However, since n = 8 000 is large and p = 0.001 is small,

$$P(X = x) \approx \frac{\lambda^x e^{-\lambda}}{x!}, \qquad x = 0, 1, 2, \ldots$$

where $\lambda = 8\,000 \times 0.001 = 8$. Thus,

$$P(X < 7) \approx \sum_{x=0}^{6} \frac{8^x e^{-8}}{x!}.$$

Using Table $A.2$ in the Appendix, we obtain $P(X<7)\approx 0.313$. The exact value is 0.313 252 073. It can be seen that the approximation is very close.

(vi) *The negative binomial distribution*

Definition 1.20

Theorem 1.18

$$P(Y \le 3) = \sum_{y=2}^{3} \binom{y-1}{1}(0.3)^2(0.7)^{y-2}$$

$$= \binom{1}{1}(0.3)^2 + \binom{2}{1}(0.3)^2(0.7) = (0.3)^2\{1 + 2(0.7)\} = 0.216.$$

Exercise 1(c)

1. A large consignment of electrical fuses contains 5% that are defective. If a sample of size 10 is picked at random from the consignment, find the probability that
   (a) no defective will be observed, (b) one defective will be observed,
   (c) at most two defectives will be observed, (d) at least one defective will be observed.

2. The probability that a drug cures a particular ailment is 0.45. If 12 people are reported sick of that ailment and given the drug, find the probability that:
   (a) all the 12 patients are cured, (b) at least 8 patients are cured,
   (c) less than 3 patients are cured.

3. A quality control officer randomly selected 10 bulbs from a large consignment of bulbs known to contain 30% defective bulbs. Find the probability that:
   (a) at least 3, (b) less than 4 of the bulbs selected are defective.

4. The probability that a person suffering from migraine headache will obtain relief with a particular drug is 0.9. Three randomly selected sufferers from migraine headache are given the drug. Find the probability that the number who obtain relief is:
   (a) exactly one, (b) 2 or less, (c) more than one, (d) exactly 3.

5. Changes in airport procedures require considerable planning. Arrival rates of aircraft is an important factor that must be taken into account. Assume that the number of small aircraft arriving at a certain airport in an hour, has the Poisson distribution with mean six.
   (a) Find the probability that three small aircraft arrive in:
      (i) one hour, (ii) thirty minutes.
   (b) What is the mean number of arrivals during a two hour period?

6. The number of customers arriving per hour at a certain automobile service facility is assumed to follow a Poisson distribution with mean $\lambda = 7$.
   (a) Calculate the probability that:
      (i) more than 10 customers will arrive in a two hour period,
      (ii) more than four customers will arrive during a one hour period.
   (b) What is the mean number of arrivals during a two hour period?

7. In a certain book, the number of errors per page has the Poisson distribution with mean 0.5.
   (a) Calculate the probability that:
      (i) a page has at least one error, (ii) two pages have three errors.
   (b) Find the mean number of errors in a book of 48 pages.

8. In a certain shop, customers arrive at the checkout counter according to a Poisson distribution at an average of 10 per hour.
   (a) During a particular one-hour period, what is the probability that
   (i) at least 3 customers arrive?, (ii) less than 5 customers arrive?
   (b) What is the probability that in a given 30 minutes, exactly two customers arrive?

9. A random variable X has the Poisson distribution with mean μ. If $P(X=2) = 0.6P(X=1)$, calculate: (a) the value of μ, (b) $P(X \geq 1)$.

10. If X is a random variable with moment generating function $M_X(t) = e^{3(e^t - 1)}$, find
    (a) $P(X=1)$, (b) $P(X \geq 2)$.

11. The random variable X has the Poisson distribution with mean μ. If $P(X=2) = 0.4P(X=1)$, calculate: (a) the value of μ, (b) $P(X \geq 2)$.

12. In a study of suicides, Gibbons et al. (1990) found that the monthly distribution of adolescent suicides in Cook Country, Illinois, between 1977 and 1987 has the Poisson distribution with mean 2.75. Find the probability that:
    (a) 3 adolescent suicides occurred in a month,
    (b) 3 or 4 adolescent suicides occurred in a month.

13. In the study of a certain aquatic organism, a large number of samples were taken from a pond, and the number of organisms in each sample was counted. The average number of organisms per sample was found to be two. Assuming that the number of organisms follows a Poisson distribution, find the probability that the next sample taken will contain:
    (a) one or fewer organisms, (b) more than five organisms.

14. A large shipment of disposable flashlights contains 1% that are defective. Use the Poisson approximation to the binomial distribution to find the probability that among 200 flashlights randomly selected from the shipment:
    (a) exactly 3 will be defective, (b) at most 2 will be defective,
    (c) at least 3 will be defective.

15. The number of flaws in a thin copper wire has the Poisson distribution with a mean of 2.3 flaws per millimetre. Find the probability of

(a) exactly 2 flaws in 1 millimetre of wire, (b) 10 flaws in 5 millimetres of wire,
(c) at least one flaw in 2 millimetres of wire.

16. Suppose that, X, the number of customers that enter a bank in an hour is a Poisson random variable, and suppose that $P(X = 0) = 0.05$. Find $E(X)$.

17. A manufacturer of a consumer electronics product expects 2% of units to fail during the warranty period. A sample of 500 independent units is tracked for warranty performance.
    (a) Find the probability that:
        (i) none fails during the warranty period,
        (ii) more than two units fail during the warranty period.
    (b) What is the expected number of failures during the warranty period?

18. If X is a random variable with characteristic function $\phi_X(t) = e^{2(e^{it}-1)}$, find:
    (a) $P(X = 0)$, (b) $P(X \geq 1)$.

19. Suppose that, for a certain couple, the probability of a male birth is 0.51 and that of a female birth is 0.49. If the couple decides to continue having children until they have one daughter, what is the probability that they will have 5 children?

20. Consider the experiment of rolling a fair six-sided die until a two occurs. Let X denote the required number of rolls.
    (a) Write down the probability mass function of X.
    (b) Find the probability that:
        (i) 5 rolls of the die are required, (ii) more than 5 rolls of the die are required.

21. The probability that a sample of air contains a rare molecule is 0.01. If it is assumed that the samples are independent with respect to the presence of the rare molecule, what is the probability that exactly 125 samples need to be analyzed before the rare molecule is detected?

22. Assume that each of your calls to a popular radio station has a probability of 0.02 of connecting, that is, of not obtaining a busy signal. Assume that your calls are independent. What is the probability that your first call that connects is your tenth call?

23. An oil prospector drills a succession of holes in a given area to find a productive well. The probability that he is successful on a given trial is 0.25.
    (a) Find the probability that the fifth hole drilled is the first that yields a productive well.
    (b) If the prospector can only afford to drill at most 8 wells, what is the probability that he fails to find a productive well?

1.6.4 Continuous random variables

Corresponding to every continuous random variable X, there is a function f, called the probability density function (p.d.f.) of X such that

(a) $f_X(x) \geq 0,$ (b) $\int_{-\infty}^{\infty} f_X(x)dx = 1,$ (c) $P(a \leq X \leq b) = \int_a^b f_X(x)dx.$

For a complete characterization of a continuous random variable, it is necessary and sufficient to know the p.d.f. of the random variable.

A consequence of X being a continuous random variable is that for any value in the range of X, say x

$$P(X = x) = \int_x^x f_X(t)dt = 0. \quad \ldots\ldots(1.6.4)$$

As an immediate consequence of Equation (1.6.4), if X is a continuous random variable, then for any numbers a and b, with $a \leq b$,

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b). \quad \ldots\ldots(1.6.5)$$

That is, it does not matter whether we include an endpoint of the interval or not. This is not true though, when X is discrete.

Example 1.22

A random variable X has p.d.f. given by

$$f_X(x) = \begin{cases} 1/10, & 0 < x < 10, \\ 0, & \text{elsewhere.} \end{cases}$$

Find: (a) $P(X > 8 | X > 5)$, (b) $P(X > 7 | X < 9)$.

Solution

(a) $$P(X > 8 | X > 5) = \frac{P(X > 8,\ X > 5)}{P(X > 5)} = \frac{P(X > 8)}{P(X > 5)}$$

$$= \left(\int_8^{10} \tfrac{1}{10}dx\right) \Big/ \left(\int_5^{10} \tfrac{1}{10}dx\right)$$

$$= \left[\tfrac{1}{10}x\right]_8^{10} \Big/ \left[\tfrac{1}{10}x\right]_5^{10} = (10-8)/(10-5) = \tfrac{2}{5}.$$

(b) $$P(X > 7 | X < 9) = \frac{P(X > 7,\ X < 9)}{P(X < 9)} = \frac{P(7 < X < 9)}{P(X < 9)}$$

$$= \left(\int_7^{9} \tfrac{1}{10}dx\right) \Big/ \left(\int_0^{9} \tfrac{1}{10}dx\right)$$

$$= \left[\tfrac{1}{10}x\right]_7^9 \Big/ \left[\tfrac{1}{10}x\right]_0^9 = (9-7)/(9-0) = \tfrac{2}{9}.$$

Definition 1.21

The p.d.f. of the continuous random variable X is given by

$$f_X(x) = \frac{d}{dx}F_X(x) \quad \ldots\ldots(1.6.6)$$

It can be seen that

$$F(x) = \int_{-\infty}^{x} f(t)dt. \quad \ldots\ldots(1.6.7)$$

Exercise 1(d)

1. Show that the following functions are probability density functions for some value of c and determine c.

   (a) $f_X(x) = \begin{cases} ce^{-4x}, & x \geq 0, \\ 0, & \text{elsewhere.} \end{cases}$ (b) $f_X(x) = \begin{cases} cx^2, & -1 < x < 10, \\ 0, & \text{elsewhere.} \end{cases}$

   (c) $f_X(x) = \begin{cases} c(1+2x), & 0 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$ (d) $f_X(x) = \begin{cases} \frac{1}{2}e^{-cx}, & x \geq 0, \\ 0, & \text{elsewhere.} \end{cases}$

2. Suppose that in a certain region, the daily rainfall (in inches) is a continuous random variable X with p.d.f. $f(x)$ given by

$$f_X(x) = \begin{cases} \frac{3}{4}(2x - x^2), & 0 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

   Find the probability that on a given day in this region, the rainfall is

   (a) not more than 1 inch, (b) greater than 1.5 inches,

   (c) equal to 1 inch, (d) less than 1 inch.

3. Let X be a continuous random variable with p.d.f.

$$f_X(x) = \begin{cases} -4x, & -0.5 < x < 0.0, \\ 4x, & 0.0 < x < 0.5, \\ 0, & \text{elsewhere.} \end{cases}$$

   Find: (a) $P(X \leq -0.3)$, (b) $P(X \leq 0.3)$, (c) $P(-0.2 \leq X \leq 0.2)$.

4. The pressure (measured in kg/cm^2) at a certain valve, is a random variable X whose p.d.f. is

$$f_X(x) = \begin{cases} \frac{6}{27}(3x - x^2), & 0 < x < 3, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the probability that the pressure at this valve is
(a) less than 2 kg/cm^2, (b) greater than 2 kg/cm^2, (c) between 1.5 and 2.5 kg/cm^2.

5. Let X denote the length in minutes of a long-distance telephone conversation. Assume that the p.d.f. of X is given by

$$f_X(x) = \tfrac{1}{10} e^{-x/10}, \quad x > 0.$$

(a) Verify that f is a p.d.f. of a continuous random variable.
(b) Find the probability that a randomly selected call will last:
(i) at most 7 minutes, (ii) at least 7 minutes, (iii) exactly 7 minutes.

6. A continuous random variable X has the p.d.f.

$$f_X(x) = \begin{cases} \frac{2}{27}(1 + x), & 2 < x < 5, \\ 0, & \text{elsewhere.} \end{cases}$$

Find: (a) $P(X < 4)$, (b) $P(3 < X < 4)$.

7. The proportion of people who respond to a certain mail-order solicitation is a continuous random variable X with p.d.f.

$$f_X(x) = \begin{cases} \frac{2}{5}(x + 2), & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find $P(0 < X < 0.8)$.
(b) Find the probability that more than $\frac{1}{4}$ but fewer than $\frac{1}{2}$ of the people contacted will respond to this type of solicitation.

8. A continuous random variable X has a p.d.f. given by

$$f_X(x) = \tfrac{1}{2}, \quad 1 \le x \le 3.$$

(a) Show that the area under the curve is equal to 1.
(b) Find: (i) $P(2 < X < 2.5)$, (ii) $P(X \le 1.6)$.

9. The p.d.f. of the length X millimetres of a hinge for fastening a door is

$$f_X(x) = 1.25, \quad 74.6 < x < 75.4.$$

(a) Find $P(X < 74.8)$.

(b) If the specifications for this process are from 74.7 to 75.3 millimetres, what proportion of hinges meet specifications?

10. The p.d.f. of the length Y metres, of a metal rod is

$$f_Y(y) = 2, \qquad 2.3 < y < 2.8.$$

If the specifications for this process are from 2.25 to 2.75 metres, what proportion of the bars fail to meet the specifications?

11. The p.d.f. of the time X seconds, required to complete an assembly operation is

$$f_X(x) = 0.1, \qquad 30 < x < 40.$$

(a) Determine the proportion of assemblies that require more than 35 seconds to complete.

(b) What time is exceeded by 90% of the assemblies?

12. Which of the following functions are probability density functions?

(a) $f_X(x) = \begin{cases} x, & -0.5 < x < 0.5, \\ 0, & \text{elsewhere.} \end{cases}$

(b) $g_X(x) = \begin{cases} \frac{1}{2}, & -1 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$

(c) $f_X(x) = \begin{cases} 1/x, & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$

(d) $h_X(x) = \begin{cases} \frac{1}{3}, & 0 < x < 1, \\ \frac{2}{3}, & 2 < x < 3. \end{cases}$

13. The random variable X has the p.d.f.

$$f_X(x) = \begin{cases} \frac{1}{30}, & 0 < x < 30, \\ 0, & \text{elsewhere.} \end{cases}$$

Find: (a) $P(X > 25 \mid X > 15)$, (b) $P(X < 20 \mid X > 15)$,

(c) $P(X > 15 \mid X < 22)$, (d) $P(X < 13 \mid X < 18)$.

1.6.5 Important continuous random variables

(i) *The continuous uniform distribution*

Definition 1.22

A random variable X has the continuous uniform distribution over the interval (a, b) if its p.d.f. is given by

$$f_X(x) = \begin{cases} 1/(b-a), & a \le x \le b, \\ 0, & \text{elsewhere.} \end{cases}$$

(ii) *The exponential distribution*

Definition 1.23

The random variable X is said to have the ***exponential distribution*** with parameters θ and β, if its probability density function is given by

$$f_X(x) = \begin{cases} \frac{1}{\beta}\exp\left(-\frac{x-\theta}{\beta}\right), & x \geq \theta,\ \beta > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

Very often, it is reasonable to take $\theta = 0$. This gives the ***one parameter exponential*** distribution. The exponential distribution is fundamental in the theory of the Poisson process due to its memoryless property and its relation with the Poisson distribution (see Chapter 7).

Theorem 1.19

If X has the exponential distribution with p.d.f.

$$f_X(x) = \begin{cases} \frac{1}{\beta}e^{-x/\beta}, & x \geq 0, \\ 0, & \text{elsewhere,} \end{cases}$$

then: (a) $E(X) = \beta$, (b) $V(X) = \beta^2$.

Example 1.23

Suppose the time it takes university graduates to get employment is exponentially distributed with mean 2 years. Find the probability that Yaw, who just completed the university, will get employment (a) not earlier than 3 years, (b) between 1 and 2 years.

Solution

Let X years denote the time it takes university graduates to get employment. The p.d.f. of X is

$$f_X(x) = \begin{cases} \frac{1}{2}e^{-x/2}, & x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) We are required to find $P(X \geq 3)$.

$$P(X \geq 3) = \int_3^\infty \frac{1}{2}e^{-x/2}dx = \left[-e^{-x/2}\right]_3^\infty = e^{-3/2} = 0.2231.$$

(b) We are required to find $P(1 < X < 2)$. Now,

$$P(1 < X < 2) = \int_1^2 \frac{1}{2} e^{-x/2} dx = \left[-e^{-x/2}\right]_1^2 = e^{-1/2} - e^{-1} = 0.2387.$$

(iii) *The gamma distribution*

Definition 1.24

The function Γ defined by

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx, \qquad a > 0,$$

is called the gamma function.

For any positive integer n,

$$\Gamma(n) = (n-1)!.$$

Definition 1.25

The continuous random variable X, is said to have the gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ if its p.d.f. is given by

$$f_X(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, & x \geq 0, \\ 0, & \text{elsewhere,} \end{cases}$$

where β is the scale parameter and α is the shape parameter of the distribution.

(iv) *The Erlang distribution*

Definition 1.26

The random variable X has the Erlang distribution with parameters n and β, if the p.d.f. of X is given by

$$f_X(x) = \frac{x^{n-1} e^{-x/\beta}}{\beta^n (n-1)!}, \qquad x \geq 0. \quad \text{...}(1.6.8)$$

It can be seen that the Erlang distribution is a special case of the gamma distribution.

Theorem 1.20

Let $X_1, X_2, \ldots, X_n$ be independent exponential random variables, each with mean β. Then, $Y = \sum_{i=1}^n X_i$ has the Erlang distribution with parameters n and β.

(v) *The normal distribution*

We now consider the most important distribution in statistics – the normal distribution. Many mathematicians figure prominently in the history of the normal distribution, including [6]Gauss.

Definition 1.27

The random variable X has the normal distribution with mean μ and variance σ^2 if its p.d.f. is given by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty. \quad \text{(1.6.9)}$$

The notation, X is $N(\mu,\ \sigma^2)$, means that the random variable X has the normal distribution with mean μ and variance σ^2. If $\mu = 0$ and $\sigma = 1$, then X is called the ***standard normal distribution***. Its distribution function is denoted by Φ. Thus,

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\, dt. \quad \text{(1.6.10)}$$

Values of this function are given in Table A.3, in the Appendix. By symmetry,

$$\Phi(-x) = 1 - \Phi(x). \quad \text{(1.6.11)}$$

Theorem 1.21

If we define $Y = aX + b$, where X is $N(\mu,\ \sigma^2)$, then we find that Y is $N(a\mu + b,\ a^2\sigma^2)$.

In particular, $Z = \dfrac{X-\mu}{\sigma}$ is $N(0,\ 1)$.

Theorem 1.22 (*The normal approximation to the binomial distribution*)

Let X be a binomial random variable with parameters n and p. For large n,

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \quad \text{(1.6.12)}$$

is approximately $N(0, 1)$.

[6]Carl Friedrich Gauss, 1777 – 1855, was born and died in Germany. He carried out numerous works in astronomy and physics, in addition to his important mathematical discoveries. He was interested, in particular, in algebra and geometry. He introduced the law of errors, that now bears his name, as a model for the errors in astronomical observations.

Theorem 1.23 (*The distribution of a linear combination of independent normally distributed random variables*)

Let X_i be $N(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots, n$. If $X_1, X_2, ..., X_n$ are independent and $c_1, c_2, ..., c_n$ are constants, then $Y = \sum_{i=1}^{n} c_i X_i$ is $N\left(\sum_{i=1}^{n} c_i\mu_i, \sum_{i=1}^{n} c_i^2\sigma_i^2\right)$.

Theorem 1.24

Let $\bar{X}$ denote the mean of a random sample of size n from a population which is $N(\mu, \sigma^2)$. Then, $\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ is $N(0, 1)$.

Example 1.24

Suppose that the length of life, measured in hours, of any one of a certain type of fuse, is an exponential random variable with mean 100 hours. What is the probability that the combined length of life of two of these fuses will be between 180 and 220 hours?

Solution

Let X_1 denote the length of life of the first fuse and X_2 the length of life of the second fuse. Then, $Y = X_1 + X_2$ is the combined length of life of the two fuses. By Theorem 1.20, Y has the Erlang distribution with p.d.f.

$$f_Y(y) = \begin{cases} \frac{1}{100^2} y e^{-y/100}, & y \geq 0, \\ 0, & y < 0. \end{cases}$$

We are required to find $P(180 < Y < 220)$.

$$\begin{aligned} P(180 < Y < 220) &= \int_{180}^{220} \frac{1}{100^2} y e^{-y/100} dy \\ &= \left[-\frac{1}{100} y e^{-y/100}\right]_{180}^{220} + \int_{180}^{220} \frac{1}{100} e^{-y/100} dy \\ &= \frac{1}{100}\left[-220e^{-2.2} + 180e^{-1.8}\right] + \left[-e^{-y/100}\right]_{180}^{220} \\ &= 2.8e^{-1.8} - 3.2e^{-2.2} \\ &= 0.1083. \end{aligned}$$

Exercise 1(e)

1. A bus arrives every 10 minutes at a bus stop. It is assumed that the waiting time for a particular individual is a random variable with a continuous uniform distribution.
   (a) Find the probability that an individual waits more than 7 minutes.
   (b) Find the probability that an individual waits between 2 and 7 minutes.

2. Suppose X has a continuous uniform distribution over the interval (–1, 1).
   (a) Find the mean and the variance of X.
   (b) Determine the value of a such that $P(-a < X < a) = 0.90$.

3. The time in seconds required to complete an assembly operation is uniformly distributed over the interval (30, 40).
   (a) Determine the proportion of assemblies that require more than 35 seconds to complete.
   (b) What time is exceeded by 90% of the assemblies?

4. Suppose that the annual rainfall in a certain region is a continuous uniform random variable with values ranging from 12 to 15 centimetres. Find the probability that in a given year, the region's rainfall will be less than 13 centimetres.

5. Suppose X has the exponential distribution with mean 10. Determine the following:
   (a) $P(X > 10)$, (b) $P(X > 20)$, (c) $P(X > 30)$,
   (d) the value of x such that $P(X < x) = 0.95$.

6. Suppose the counts per minute recorded by a geiger counter have the Poisson distribution with mean 2.
   (a) What is the probability that there are no counts in a 30-second interval?
   (b) What is the probability that the first count occurs in less than 10 seconds?
   (c) What is the mean time between counts?
   (d) Determine x, such that the probability that at least one count occurs before the time x minutes is 0.95.

7. The time between arrivals of electronic messages at your computer is exponentially distributed with a mean of 2 hours.
   (a) What is the probability that you do not receive a message in a two-hour period?
   (b) If you have not had a message in the last 4 hours, what is the probability that you receive a message in the next 2 hours?
   (c) What is the expected time between your fifth and sixth messages?

8. Suppose that the operation lifetime of a certain type of battery is an exponential random variable with mean θ (measured in years). Find the probability that
   (a) a battery of this type will have an operating lifetime of over 4 years.
   (b) at least one of 5 batteries of this type will have operating lifetime of over 4 years.

9. The lifetime of a mechanical assembly in a vibration test is exponentially distributed with a mean of 400 hours.
   (a) What is the probability that an assembly on test fails in less than 100 hours?
   (b) What is the probability that an assembly operates for more than 500 hours before failure?
   (c) If an assembly has been on test for 400 hours without a failure, what is the probability of a failure in the next 100 hours?
   (d) If 10 assemblies are placed on test, what is the probability that at least one fails in less than 100 hours? Assume that the assemblies fail independently.

10. Find: (a) $\int_0^\infty x^2 e^{-x} dx$, (b) $\int_0^\infty x^7 e^{-2x} dx$, (c) $\int_0^\infty x^4 e^{-6x} dx$, (d) $\int_0^\infty x^5 e^{-3x} dx$.

11. If X has the gamma distribution with parameters $\alpha = 2$ and $\beta = 4$, find:
    (a) $E(X)$, (b) $V(X)$, (b) $P(X \le 4)$.

12. In a certain city, the daily consumption of water (in millions of litres) follows a gamma distribution with $\alpha = 2$ and $\beta = 3$. If the daily capacity of that city is 9 million litres of water, what is the probability that on any given day the water supply is inadequate?

13. If the total cholesterol values for a certain population are approximately normally distributed with a mean of 200 mg/100 ml and a standard deviation of 20 mg/100 ml, find the probability that an individual picked at random from this population will have a cholesterol value:
    (a) between 180 and 200 mg/100 ml, (b) greater than 225 mg/100 ml,
    (c) less than 150 mg/100 ml, (d) between 190 and 210 mg/100 ml.

14. The IQs of individuals admitted to a state school for mentally retarded are approximately $N(60, 100)$.
    (a) Find the proportion of individuals with IQs greater than 75.
    (b) Find the probability than an individual picked at random will have an IQ between 55 and 75.

15. If X is $N(8, \sigma^2)$, find:
    (a) $P(X > 8)$, (b) $P(8-\sigma < X < 8+\sigma)$, (c) $P(8-2\sigma < X < 8+3\sigma)$.

16. The heights of students of St. Andrew School are normally distributed with mean 160 cm and standard deviation 8 cm.
    (a) What proportion of students have heights between 152 cm and 176 cm?
    (b) If 25% of the students are more than h cm tall, find the value of h.

17. An electrical firm manufactures light bulbs that have a life, before burn-out, that is normally distributed with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a bulb burns between 778 and 834 hours.

18. The probability that a patient recovers from a delicate heart operation is 0.9. Of the next 100 patients having this operation, what is the probability that
    (a) between 84 and 95 inclusive survive? (b) fewer than 86 survive?

1.6.6 Transformations

If X is a random variable, then any transformation $Y = g(X)$, where g is a real-valued function defined on R, is also a random variable.

Theorem 1.25

Let X be a continuous random variable with p.d.f. $f_X(x)$ and let $Y = g(X)$, where $y = g(x)$ is a one-to-one differentiable function with inverse $x = g^{-1}(y)$. Then, the p.d.f. of Y is given by

$$f_Y(y) = f_X\left\{g^{-1}(y)\right\}\left|\frac{d}{dy}\left(g^{-1}(y)\right)\right|.$$

Theorem 1.26 (*The probability integral transformation*)

If X is a continuous random variable with distribution function $F_X(x)$, then $Y = F_X(X)$ has the continuous uniform distribution over the interval (0, 1).

The transformation $Y = F_X(X)$ is called the ***probability integral transformation***.

Example 1.25

The random variable X has the continuous uniform distribution over the interval (0, 1).

(a) Find the p.d.f. of $Y = -2\ln X$.

(b) If $X_1, X_2, \ldots, X_n$ is a random sample of size n of X, what is the distribution of $W = -\sum_{i=1}^{n} 2\ln X_i$? Give reasons for your answer.

Solution

(a) The p.d.f. of X is given by

$$f_X(x) = 1, \qquad 0 < x \le 1.$$

$y = -2\ln x \Rightarrow x = e^{-\frac{1}{2}y}$ and $0 < x \le 1 \Rightarrow y \ge 0.$

Moreover, y is a monotone (decreasing) function of x. Hence, the p.d.f. of Y is given by

$$f_Y(y) = f_X\left(e^{-\frac{1}{2}y}\right)\left|\frac{d}{dy}e^{-\frac{1}{2}y}\right| = \left|-\frac{1}{2}e^{-\frac{1}{2}y}\right| = \frac{1}{2}e^{-\frac{1}{2}y}, \quad y \ge 0.$$

(b) W is the sum of n independent random variables, each with the exponential distribution with mean 2. Hence, W has the Erlang distribution with p.d.f. (see Theorem 1.20 on page 33).

$$f_W(w) = \frac{w^{n-1}e^{-w/2}}{2^n(n-1)!}, \qquad w \ge 0.$$

Exercise 1(f)

1. The random variable X has the continuous uniform distribution over the interval (–1, 1).
   (a) Find the distribution function and the p.d.f. of $Y = |X|$.
   (b) Find the p.d.f. of $Y = X^2$.
2. Let X be a continuous random variable with p.d.f. given by
   $$f_X(x) = \frac{1}{2}e^{-|x|}, \qquad -\infty < x < \infty.$$
   Find the p.d.f. of (a) $Y = X^2$, (b) $W = |X|$.
3. Given that the p.d.f. of X is $f_X(x) = \theta x^{\theta-1}$, $0 < x < 1$, find the p.d.f. of $Y = -\ln X$.
4. The random variable X has the p.d.f. given by
   $$f_X(x) = 2x, \qquad 0 < x < 1.$$
   Find the distribution functions of (a) $Y = X^2$, (b) $W = X^3$.

5. The random variable X has the continuous uniform distribution over the interval (0, 1). Find the p.d.f. of (a) $Y = -3\ln X$, (b) $W = -4\ln X$.

6. The random variable X has the p.d.f. given by

$$f_X(x) = 2xe^{-x^2}, \qquad x \geq 0.$$

Find the p.d.f. of $Y = X^2$.

7. The random variable X has the p.d.f. given by

$$f_X(x) = \theta / x^{\theta+1}, \qquad x \geq 1.$$

(a) Show that $Y = \ln X$ has the exponential distribution and find $E(Y^3)$.

(b) If $X_1, X_2, ..., X_n$ is a random sample of size n of X, find the p.d.f. of $W = \sum_{i=1}^{n} \ln X_i$.

8. The random variable X has the p.d.f. given by

$$f_X(x) = \tfrac{1}{9}x^2, \qquad 0 < x < 3.$$

(a) If $Y = \frac{1}{27}X^3$, find the p.d.f. of Y. (b) If $U = X^2$, find the p.d.f. of U.

9. The random variable X has the p.d.f. given by

$$f_X(x) = 2(1-x), \qquad 0 < x < 1.$$

Find the p.d.f. of $Y = \ln(1-X)$.

10. The random variable X has the exponential distribution with mean θ

12. Find the transformation $Y = H(X)$ such that if X has the p.d.f.

$$f_X(x) = 4x^3, \qquad 0 < x < 1,$$

then Y has the p.d.f.

$$g_Y(y) = e^y \Big/ \left(1 + e^y\right)^2, \qquad -\infty < y < \infty.$$

13. The random variable U has the continuous uniform distribution on the interval (0, 1). Find a transformation $Y = G(U)$ such that the distribution of Y has the p.d.f.

$$f_Y(y) = e^{-y} \Big/ \left(1 + e^{-y}\right)^2, \qquad -\infty < y < \infty.$$

14. The random variable U has the continuous uniform distribution over the interval (0, 1). Find a transformation $Y = H(U)$ such that the distribution of Y is:

(a) exponential with p.d.f. $\lambda e^{-\lambda y}, \quad y \geq 0, \ \lambda > 0.$

(b) logistic with p.d.f. $e^y \Big/ \left(1 + e^y\right)^2, \quad -\infty < y < \infty.$

15. The random variable U has the continuous uniform distribution on the interval (0, 1). Find a transformation $Y = G(U)$ such that the distribution of Y has the p.d.f.

$$f_Y(y) = cy^{c-1}\exp(-y^c), \qquad (y > 0, \ c > 0).$$

16. If X has the continuous uniform distribution over the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, find the distribution of $Y = \tan X$.

17. If X has the continuous uniform distribution over the interval (0, 1), find the distribution of $Y = 1/X$.

18. (a) If U has the continuous uniform distribution over the interval (0, 1), find the p.d.f. of $X = -\ln U$.

(b) Let $U_1, U_2, \ldots, U_n$ be a random sample from the continuous uniform distribution over the interval (0, 1) and let $Y = \sum_{j=1}^{n} (-\ln U_j)$ and $W = (U_1 U_2 \ldots U_n)^{1/n}$. Find the p.d.f. of (i) Y, (ii) $T = -n\ln W$, (iii) W.

19. Suppose X has the Pareto distribution with p.d.f.

Solution

$$\text{Expected amount of food} = E\big(g(X)\big) = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx$$

$$= \int_0^2 \tfrac{3}{4} x^2 (2x - x^2)\, dx = \tfrac{3}{4}\int_0^2 (2x^3 - x^4)\, dx$$

$$= \tfrac{3}{4}\left[\tfrac{1}{2}x^4 - \tfrac{1}{5}x^5\right]_0^2 = \tfrac{3}{4}\left(8 - \tfrac{32}{5}\right) = 6 - \tfrac{24}{5} = 1.2 \text{ mg.}$$

1.6.8 Characteristic functions

Definition 1.29 (*Characteristic function*)

The characteristic function of a random variable X, denoted by $\phi_X(t)$, is defined for all real t, by $\phi_X(t) = E\left(e^{itX}\right)$, where $i^2 = -1$.

If X is a continuous random variable, then $\phi_X(t)$ is the [7]Fourier transform of the density function $f_X(x)$:

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x)\, dx. \qquad \text{(1.6.13)}$$

We can invert this Fourier transform and obtain

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi_X(t)\, dt. \qquad \text{(1.6.14)}$$

This result is called the ***inversion formula***. It can be used to find the p.d.f. of a random variable if we know the characteristic function of the random variable. Since the Fourier transform of a function is unique, the function $\phi_X(t)$ characterizes entirely the random variable X. For instance, only the standard normal distribution has the characteristic function $\phi_Z(t) = e^{-\frac{1}{2}t^2}$.

Example 1.27

Find the p.d.f. of the random variable X whose characteristic function is given by $\phi_X(t) = e^{-|t|}$.

---

[7] Joseph (Baron) Fourier, 1768 – 1830, was born and died in France. He taught at the Collège de France and at the École Polytechnique. In his main work, the Théorie Analytique de la Chaleur, published in 1822, he made wide use of the series which now bears his name, but which he did not invent.

Solution

The p.d.f. of X is given by

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}e^{-|t|}dt = \frac{1}{2\pi}\int_{-\infty}^{0} e^{-itx}e^{t}dt + \frac{1}{2\pi}\int_{0}^{\infty} e^{-itx}e^{-t}dt$$

$$= \frac{1}{2\pi}\int_{-\infty}^{0} e^{t(1-ix)}dt + \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-t(1+ix)}dt$$

$$= \frac{1}{2\pi}\left[\frac{e^{t(1-ix)}}{1-ix}\right]_{-\infty}^{0} + \frac{1}{2\pi}\left[\frac{-e^{-t(1+ix)}}{1+ix}\right]_{0}^{\infty} = \frac{1}{2\pi}\left(\frac{1}{1-ix}\right) + \frac{1}{2\pi}\left(\frac{1}{1+ix}\right)$$

$$= \frac{1}{2\pi}\left(\frac{1+ix+1-ix}{1-i^2x^2}\right) = \frac{1}{\pi(1+x^2)}, \qquad -\infty < x < \infty.$$

Remark

The random variable X in Example 1.27 has the standard [8]Caushy distribution.

We can also use the characteristic function to obtain the moments of order n of a random variable. The result is given in the following theorem.

Theorem 1.27

If the n^{th} moment of X exists, then the n^{th} derivative $\phi_X^{(n)}(t)$, of $\phi_X(t)$ exists, and

$$\phi_X^{(k)}(0) = i^k E\left(X^k\right), \qquad k = 1, 2, ..., n. \quad \text{...(1.6.15)}$$

Many authors prefer to work with the following function, which, as its name indicates, also enables us to calculate the moments of a random variable.

Definition 1.30 (*Moment generating function*)

The moment generating function of a random variable X, denoted by $M_X(t)$, is defined by

$$M_X(t) = E\left(e^{tX}\right) = \begin{cases} \sum_{i=0}^{\infty} e^{tx_i} p_X(x_i), & \text{if } X \text{ is discrete.} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx, & \text{if } X \text{ is continuous.} \end{cases}$$

[8]Augustin Louis Cauchy, 1789 – 1857, was born and died in France. He is considered the father of mathematical analysis and the inventor of the theory of functions of a complex variable.

Remarks

(i) When X is a continuous and non-negative random variable, $M_X(t)$ is the [9]Laplace transform of $f_X(x)$.

(ii) Corresponding to Theorem 1.27, we obtain the following theorem.

Theorem 1.28

$$E\left(X^n\right) = \frac{d^n}{dt^n} M_X(t)\Big|_{t=0} \quad \text{for } n = 1,\ 2,\ \ldots \qquad (1.6.16)$$

Example 1.28

If X has the Poisson distribution with parameter λ, then

$$M_X(t) = \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(e^t\lambda\right)^k}{k!} = e^{-\lambda} \exp\left(e^t\lambda\right).$$

We deduce from this result that:

$$E(X) = M_X'(0) = \lambda \quad \text{and} \quad E(X^2) = M_X''(0) = \lambda^2 + \lambda.$$

It follows that

$$V(X) = \lambda^2 + \lambda - (\lambda)^2 = \lambda = E(X).$$

Note that to obtain $E(X^2)$, we can proceed as follows:

$$E\left(X^2\right) = \sum_{k=0}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} = e^{-\lambda}\lambda \sum_{k=1}^{\infty} \frac{d}{d\lambda} \frac{\lambda^k}{(k-1)!}$$

$$= e^{-\lambda}\lambda \frac{d}{d\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = e^{-\lambda}\lambda \frac{d}{d\lambda}\left(\lambda e^{\lambda}\right) = \lambda + \lambda^2.$$

It is clear that it is easier to differentiate twice the function $e^{-\lambda}\exp(e^t\lambda)$ than to evaluate the above infinite sum.

If we do not know the distribution of the random variable X, we can use the following inequalities to obtain bounds for the probability of certain events.

---

[9]Pierre Simon (Marquis) Laplace, 1749 – 1827, was born and died in France. In addition to being a mathematician and astronomer, he was also a minister and a senator. He participated in the organization of École Polytechnique of Paris. His main works were on astronomy and on the calculus of probabilities: the Traité de Mécanique des Probabilités, whose first edition appeared in 1812. Many mathematical formulae bear his name.

Theorem 1.29 ([10]***Markov's inequality***)

Theorem 1.30 ([11]***Chebyshev's inequality***)

1. Consider the random variable X with probability mass function given by

   $$p_X(x) = \frac{1}{5}(x - 3)^2, \qquad x = 3, 4, 5.$$

   (a) Find

Find

(a) Show that f is a probability mass function of a discrete random variable.

(b) Find $E(X)$ and $E(X^2)$ from the definitions of these terms.

(c) Find $M_X(t)$ and use it to verify your answers to part (b).

(d) Find $V(X)$.

4. Let X be a discrete random variable with three possible values and probabilities as follows: $p_X(0)=\frac{1}{4}$, $p_X(1)=\frac{1}{2}$ and $p_X(2)=\frac{1}{4}$.

(a) Find $E(X)$ and $E(X^2)$ from the definitions of these terms.

(b) Find $M_X(t)$ and use it to verify your answers to part (a).

5. The random variable Y has the p.d.f.

$$f_Y(y)=\begin{cases}2y, & 0\le y\le 1,\\ 0, & \text{elsewhere.}\end{cases}$$

(a) Find the moment generating function of Y.
(b) Using the moment generating function, find $E(Y)$ and $V(Y)$.

6. The random variable X has the p.d.f.

$$f_X(x)=\begin{cases}xe^{-x}, & x\ge 0,\\ 0, & \text{elsewhere.}\end{cases}$$

(a) Find the moment generating function of X.
(b) Find the mean and variance of X from its moment generating function.

7. X and Y are independent random variables with respective moment generating functions

$$M_X(t)=e^{3(e^t-1)} \quad\text{and}\quad M_Y(t)=e^{2(e^t-1)}.$$

Find (a) the moment generating function of $X + Y$, (b) $E(X+Y)$.

8. Consider the random variable X with probability mass function

$$p_X(x)=\tfrac{1}{5}(x-2)^2,\quad x=2,\ 3,\ 4.$$

(a) Find $E(X)$. (b) Find the characteristic function of X and use it to find $E(X)$.

(c) Find $E(X^2)$. (b) Use the characteristic function of X to find $E(X^2)$.

9. A discrete random variable has characteristic function $\phi_X(t)=e^{3(e^{it}-1)}$. Find $E(X)$ and $V(X)$.

10. Let X be a discrete random variable with three possible values and probabilities as follows: $p_X(0)=\frac{1}{3}$, $p_X(1)=\frac{1}{2}$, $p_X(2)=\frac{1}{6}$.

(a) Find $E(X)$ and $E(X^2)$ from the definitions of these terms.

(b) Find $\phi_X(t)$ and use it to verify your answers to part (a).

11. The random variable X has p.d.f.

$$f_X(x) = \frac{1}{2}e^{-x/2}, \quad x \geq 0.$$

(a) Find the characteristic function of X. (b) Use $\phi_X(t)$ to find $E(X)$ and $V(X)$.

12. Find the p.d.f. of the random variable X whose characteristic function is given by $\phi_X(t) = e^{-\frac{1}{2}t^2}$.

13. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n of X. If $\phi_X(t) = (1-2it)^{-\frac{1}{2}v}$, find the characteristic function of

(a) $Y_n = \sum_{j=1}^{n} X_j$, (b) $W_n = (Y_n - n)\big/\sqrt{2n}$.

14. Let X, Y and Z be independent binomial random variables with parameters (n_1, p), (n_2, p) and (n_3, p), respectively.

(a) Find the moment generating function of $W = X + Y + Z$.

(b) Find the characteristic function of $W = X + Y + Z$.

(c) What is the distribution of W? Give reasons for your answer.

15. The random variable X has the moment generating function given by

$$M_X(t) = (2 + 3e^t)^6 / 5^6.$$

(a) What is the distribution of X? Give reasons for your answer. (b) Find $P(X \geq 1)$.

16. A random variable Y has the characteristic function given by

$$\phi_Y(t) = \left(e^{it} + 4\right)^7 \Big/ 5^7.$$

(a) What is the distribution of Y? Give reasons for your answer.

(b) Find: (i) $P(Y = 2)$, (ii) $P(Y \geq 1)$.

1.7 Random vectors

Definition 1.31

An n-dimensional random vector is a function $X = (X_1, X_2, \ldots, X_n)$ that associates a vector $(X_1(s), X_2(s), \ldots, X_n(s))$ of real numbers with each element s of a sample space S of a random experiment E. Each component X_k of the vector is a random variable.

1.7.1 Two dimensional random vectors

Definition 1.32

The joint distribution of the random vector (X, Y) is defined, for all points $(x, y) \in R$, by

$$F_{X,Y}(x, y) = P(\{X \le x\} \cap \{Y \le y\}) \equiv P(X \le x, Y \le y). \quad \ldots\ldots\ldots\ldots\ldots(1.7.1)$$

Properties

(i) $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0$ and $F_{X,Y}(\infty, \infty) = 1$.

(ii) $F_{X,Y}(x_1, y_1) \le F_{X,Y}(x_2, y_2)$ if $x_1 \le x_2$ and $y_1 \le y_2$.

(iii) We have

$$\lim_{\varepsilon \to 0} F_{X,Y}(x+\varepsilon, y) = \lim_{\varepsilon \to 0} F_{X,Y}(x, y+\varepsilon) = F_{X,Y}(x, y). \quad \ldots\ldots\ldots\ldots\ldots\ldots(1.7.2)$$

The marginal distribution function of X, when the function $F_{X,Y}$ is known, is obtained as follows

$$F_X(x) = P(X \le x, Y < \infty) = F_{X,Y}(x, \infty). \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1.7.3)$$

Definition 1.33

A two dimensional random vector $\boldsymbol{Z} = (X, Y)$, is of discrete type if it takes a finite or a countably infinite set of points in R^2.

Definition 1.34

The joint probability mass function of the discrete random variable (X, Y) is defined by

$$p_{X,Y}(x_j, y_k) = P(X = x_j, Y = y_k) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1.7.4)$$

for $j, k = 1, 2, \ldots$.

To obtain the marginal probability mass function of X, from $p_{X,Y}$, we make use of the total probability rule:

$$p_X(x_j) = P(X = x_j) = \sum_{k=1}^{\infty} P(\{X = x_j\} \cap \{Y = y_k\}) = \sum_{k=1}^{\infty} p_{X,Y}(x_j, y_k). \quad \ldots(1.7.5)$$

Definition 1.35

The joint probability density function of the continuous random vector $Z=(X,\ Y)$ is defined by

$$f_{X,Y}(x,\ y)\ =\ \frac{\partial^2}{\partial x\,dy}F_{X,Y}(x,\ y).\ \ldots\ldots(1.7.6)$$

Notes

1. Corresponding to the formula (1.7.6), the marginal probability density function of X can be obtained as follows:

$$f_X(x)\ =\ \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy,\ \ldots\ldots(1.7.7)$$

where we integrate in practice over all the values that Y can take when $X = x$.

2. The probability of the event $\{Z\in A\}$, can be calculated as follows:

$$P(Z\in A)\ =\ \iint\limits_{A} f_{X,Y}(x,y)\,dx\,dy.\ \ldots\ldots(1.7.8)$$

3. The distribution function of the continuous random variable (X, Y) is given by

$$F_{X,Y}(x,\ y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(u,v)\,du\,dv.\ \ldots\ldots(1.7.9)$$

In the discrete case, Equation (1.7.9) becomes

$$F_{X,Y}(x,\ y)=\sum_{x\le x_j}\ \sum_{y\le y_k} p_{X,Y}(x_j,\ y_k).\ \ldots\ldots(1.7.10)$$

Example 1.29

The random variables X and Y have a joint p.d.f. given by

$$f_{X,Y}(x,y)=\begin{cases} cxy, & 0<x<y<1,\\ 0, & \text{elsewhere.}\end{cases}$$

(a) Find the value of the constant c.

(b) Find: (i) the marginal probability density functions of X and Y, (ii) $E(Y)$.

Solution

(a) The value of c is given by the equation

$$1\ =\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,\ y)\,dx\,dy\ =\ \int_0^1\int_0^y cxy\,dx\,dy\ =\ c\int_0^1\left[\frac{x^2y}{2}\right]_0^y dy\ =\ c\int_0^1\frac{1}{2}y^3dy\ =\ \frac{c}{8}.$$

Hence, $c\ =\ 8$.

where A (respectively, B) is any event that involves only X (respectively Y). In particular, we must have

$$F_{X,Y}(x,y) = P(X \le x,\ Y \le y) = F_X(x)F_Y(y)\ \ \forall\ (x,y). \quad \text{……………..(1.7.14)}$$

2. If X and Y are independent random variables, then so are $g(X)$ and $h(Y)$, where $g(X)$ and $h(Y)$ are real-valued functions of X and Y, respectively.
3. Let $Z = X + Y$, where X and Y are two independent random variables. Then,

$$M_Z(t) = E\left(e^{tZ}\right) = E\left(e^{t(X+Y)}\right) = E\left(e^{tX}\right)E\left(e^{tY}\right) = M_X(t)M_Y(t). \quad \text{………(1.7.15)}$$

Similarly, $\phi_Z(t) = \phi_X(t)\phi_Y(t)$.

Example 1.31

Determine whether the random variables with the following joint probability density function are independent.

$$f_{X,Y}(x,y) = \begin{cases} 8xy, & 0 < x < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Solution

First method

The ranges of X and Y are dependent and so X and Y are not independent.

Second method

From Example 1.29,

$$f_X(y) = 4x(1-x^2), \qquad 0 < x < 1.$$

and $\quad f_Y(y) = 4y^3, \qquad 0 < y < 1.$

$f_X(x)\,f_Y(y) \neq f_{X,Y}(x,y)$ for all x, y and so X and Y are not independent.

1.7.3 Conditional probability distribution

Definition 1.36

If Y is a discrete random variable, then the conditional distribution function of X, given that $Y = y_k$, is defined by

$$F_{X|Y}\left(x|y_k\right) = \frac{P\left(X \le x,\ Y = y_k\right)}{P\left(Y = y_k\right)},\ \ \text{if } P\left(Y = y_k\right) > 0. \quad \text{……………………..(1.7.16)}$$

Definition 1.37

If (X, Y) is a discrete random vector, then the conditional probability mass function of X, given that $Y = y_k$, is given by

$$p_{X|Y}\left(x_j \middle| y_k\right) = \frac{p_{X,Y}\left(x_j, y_k\right)}{p_Y\left(y_k\right)} = \frac{P\left(X = x_j, Y = y_k\right)}{P\left(Y = y_k\right)}, \text{ if } P\left(Y = y_k\right) > 0 \quad \text{.....(1.7.17)}$$

The conditional probability mass functions possess the same properties as the corresponding marginal functions.

Definition 1.38

If (X, Y) is a continuous random vector and $f_Y(y) > 0$, then the conditional distribution function and the conditional density function of X, given $Y = y$, are given respectively, by

$$F_{X|Y}\left(x \middle| y\right) = \frac{\int_{-\infty}^{x} f_{X,Y}(u, y)\,du}{f_Y(y)} \text{ and } f_{X|Y}\left(x \middle| y\right) = \frac{f_{X,Y}(x, y)}{f_Y(y)}. \quad \text{.........(1.7.18)}$$

Example 1.32

The continuous random vector (X, Y) has the p.d.f.

$$f_{X,Y}(x, y) = \begin{cases} e^{-x}, & x > y > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the conditional probability density functions of X and Y.

Solution

We first find the marginal density functions of X and Y.

$$f_X(x) = \int_0^x e^{-x} dy = \left[ye^{-x}\right]_0^x = xe^{-x}, \qquad x > 0.$$

$$f_Y(y) = \int_y^\infty e^{-x} dx = \left[-e^{-x}\right]_y^\infty = e^{-y}, \qquad y > 0.$$

The conditional probability density functions of X and Y are given by

$$f_{X|Y}\left(x \middle| y\right) = \frac{f(x, y)}{f_Y(y)} = \frac{e^{-x}}{e^{-y}} = e^{-(x-y)}, \qquad x > y > 0.$$

$$f_{Y|X}\left(y \middle| x\right) = \frac{f(x, y)}{f_X(x)} = \frac{e^{-x}}{xe^{-x}} = \frac{1}{x}, \qquad x > y > 0.$$

Independent random variables

Definition 1.39

The random variables X and Y are independent if and only if the conditional distribution function, the conditional probability mass function, or the conditional density function of X, given that $Y = y$, is identical to the corresponding marginal function.

The continuous random variables X and Y are independent if their joint p.d.f. can be written as the product of 2 non-negative functions, each involving any one of the variables, and the limits are independent.

Definition 1.40

We say that the continuous random variables X and Y have a binormal (or bivariate normal) distribution with parameters $\mu_X \in R$, $\mu_Y \in R$, $\sigma_X^2 > 0$, $\sigma_Y^2 > 0$ and $\rho \in (-1,\ 1)$, and we write that (X, Y) is $N\left(\mu_X,\ \mu_Y;\ \sigma_X^2,\ \sigma_Y^2;\ \rho\right)$, if their joint density function is given by

$$f_{X,Y}(x,\ y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}e^{-\frac{1}{2}q}, \qquad -\infty < x < \infty, \quad -\infty < y < \infty, \ldots\ldots\ldots\ldots(1.7.19)$$

where $\frac{1}{1 \quad 2} \quad \frac{\quad}{X}^{2} \quad 2 \quad \frac{\quad}{X} \quad \frac{\quad}{Y} \quad \frac{\quad}{Y}^{2}$.

It is easy to show that if (X, Y) is $N\left(\mu_X,\ \mu_Y;\ \sigma_X^2,\ \sigma_Y^2;\ \rho\right)$, then X is $N\left(\mu_X,\ \sigma_X^2\right)$ and Y is $N\left(\mu_Y,\ \sigma_Y^2\right)$. Furthermore, it can be proved that the conditional distribution of X given $Y = y$ is $N\left(\mu_X + (y-\mu_Y)\rho\sigma_X/\sigma_Y,\ \sigma_X^2(1-\rho^2)\right)$. The parameter ρ is called the ***correlation coefficient*** between X and Y. It can be seen that ***if X and Y have the bivariate normal distribution with correlation coefficient ρ, then they are independent if and only if ρ = 0.***

Example 1.33

In a certain population of married couples, the mass X of a husband and the mass Y of his wife have a bivariate normal distribution with parameters $\mu_1 = 58$ kg, $\mu_2 = 53$ kg, $\sigma_1 = \sigma_2 = 2$ kg and $\rho = 0.6$. Given that the mass of the husband is 63 kg, find the probability that his wife has a mass between 52.8 kg and 59.2 kg.

Solution

We are required to find $P(52.8 < Y < 59.2 | X = 63)$. The conditional distribution of Y given $X = 63$ is normal with mean 53 + 0.6(63 – 58) = 56 kg and standard deviation $2\sqrt{1 - 0.36}$ = 1.6 kg. Hence,

$$
\begin{aligned}
P(52.8 < Y < 59.2 | X = 63) &= P\left(\frac{52.8 - 56}{1.6} < \frac{Y - 56}{1.6} < \frac{59.2 - 56}{1.6}\right) \\
&= P(-2 < Z < 2), \quad \text{where } Z \text{ is } N(0, 1) \\
&= P(Z < 2) - P(Z < -2) = 0.954.
\end{aligned}
$$

1.7.5 Conditional expectation

Definition 1.41

The conditional expectation of X, given $Y = y$, is defined by

$$
E(X|Y = y) = \begin{cases} \sum_{j=1}^{\infty} x_j p_{X,Y}(x_j|y), & \text{(discrete case)}, \\ \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx, & \text{(continuous case)}. \end{cases}
$$

The mean, $E(g(X))$, of a transformation g of a random variable X, is a real constant, while $E\{g(X)|Y = y\}$ is a function of y, where y is a particular value taken by the random variable Y. We now consider $E\{g(X)|Y\}$. It is a function of the random variable Y that takes the value $E\{g(X)|Y = y\}$ when $Y = y$. Consequently, $E\{g(X)|Y\}$ is a random variable, whose mean can be calculated. This is given by the following important theorem.

Theorem 1.31

$$E(g(X)) = E\left[E\{g(X)|Y\}\right] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1.7.21)$$

A proof of this theorem is given by Ofosu and Hesse (2011).

Notes

1. We can deduce from Theorem 1.31, that:

$$
E(X) = E\{E(X|Y)\} = \begin{cases} \sum_{k=1}^{\infty} E(X|Y = y_k)\, p_Y(y_k), & \text{(discrete case)}, \\ \int_{-\infty}^{\infty} E(X|Y = y)\, f_Y(y)\,dy, & \text{(continuous case)}. \end{cases} \quad \ldots.(1.7.22)
$$

2. We can calculate the variance of X by conditioning on another random variable Y as follows:

$$V(X) = E\left\{E\left(X^2|Y\right)\right\} - \left[E\left\{E\left(X|Y\right)\right\}\right]^2.$$

Example 1.34

The random variable X has the Poisson distribution with mean M, where M is a random variable with p.d.f.

$$f_M(m) = m^{k-1}e^{-m}/(k-1)!, \qquad m > 0.$$

Find $E(X)$.

Solution

Using Equation (1.7.21), we obtain

$$\begin{aligned} E(X) &= E\left\{E\left(X|M\right)\right\} = E(M) \\ &= \int_0^\infty m\, f_M(m)\, dm = \int_0^\infty \frac{m^k e^{-m}}{(k-1)!} dm = \int_0^\infty \frac{m^{k+1-1}e^{-m}}{(k-1)!} dm \\ &= \frac{\Gamma(k+1)}{\Gamma(k)} = \frac{k\Gamma(k)}{\Gamma(k)} = k. \end{aligned}$$

1.7.6 Estimation of a random variable using another random variable

Suppose that we wish to estimate a random variable X by using another random variable Y. It can be shown that the function $g(Y)$ that minimizes the mean-square error (MSE)

$$MSE = E\left[\left\{X - g(Y)\right\}^2\right] \quad \text{……………………………………………..(1.7.23)}$$

is $g(Y) = E\left(X|Y\right)$. If we look for a function of the form $g(Y) = aY + \beta$, we can show that the constants α and β that minimize MSE are

$$\hat{\alpha} = \frac{E(XY) - E(X)E(Y)}{V(Y)} \text{ and } \hat{\beta} = E(X) - \hat{\alpha}E(Y). \quad \text{……………………….(1.7.24)}$$

Finally, if $g(Y) = c$, we easily find that the constant c that yields the smallest MSE is $c = E(X)$.

The function $g(Y) = E\left(X|Y\right)$ is the best estimator of X, in terms of Y, while $g(Y) = \hat{\alpha}Y + \hat{\beta}$ is the best linear estimator of X, in terms of Y. If X and Y both have the normal distribution, then the two estimators are equal.

Theorem 1.31 also enables us to calculate the probability of the event $\{X \in A\}$ by conditioning on the possible values of a random variable Y. We only have to define a random variable W such that $W = 1$ if $X \in A$ and $W = 0$ if $X \notin A$, and use the fact that $E(W) = P(X \in A)$. We can then obtain the following Theorem, which is the equivalent of Theorem 1.12, the total probability rule for random variables.

Theorem 1.32

$$P(X \in A) = \begin{cases} \sum_{k=1}^{\infty} P(X \in A | Y = y_k)\, p_Y(y_k), & \text{(discrete case)}, \\ \int_{-\infty}^{\infty} P(X \in A | Y = y)\, f_Y(y) dy, & \text{(continuous case)}. \end{cases} \quad \ldots\ldots\ldots\ldots(1.7.25)$$

Example 1.35

The conditional p.d.f. of a random variable X, given that another random variable Y has taken the value y is

$$f_{X|Y}(x|y) = ye^{-xy}, \qquad x > 0,\ y > 0.$$

If the marginal p.d.f. of Y is

$$f_Y(y) = y^{m-1}e^{-y}/\Gamma(m), \qquad m > 1,\ y > 0,$$

find $P(X > d)$.

Solution

By Theorem 1.32,

$$\begin{aligned} P(X > d) &= \int_{-\infty}^{\infty} \{P(X > d | Y = y)\} f_Y(y)\, dy = \int_{-\infty}^{\infty} \left\{ \int_d^{\infty} f_{X|Y}(x|y)\, dx \right\} f_Y(y)\, dy \\ &= \int_{-\infty}^{\infty} \left\{ \int_d^{\infty} ye^{-yx} dx \right\} f_Y(y)\, dy = \int_{-\infty}^{\infty} \left\{ \left[-e^{-yx} \right]_d^{\infty} \right\} f_Y(y)\, dy \\ &= \int_{-\infty}^{\infty} e^{-yd} f_Y(y)\, dy = \int_0^{\infty} \frac{e^{-yd}\, y^{m-1}e^{-y}}{\Gamma(m)}\, dy \\ &= \frac{1}{\Gamma(m)} \int_0^{\infty} y^{m-1} e^{-y(1+d)}\, dy. \end{aligned}$$

Let $w = y(1+d)$. Then, $dy = \frac{1}{1+d} dw$, and so

$$P(X > d) = \frac{1}{\Gamma(m)} \int_0^{\infty} \left(\frac{w}{1+d} \right)^{m-1} e^{-w} \frac{1}{1+d} dw = \frac{(1+d)^{-m}}{\Gamma(m)} \int_0^{\infty} w^{m-1} e^{-w} dw$$

$$= \frac{\Gamma(m)(1+d)^{-m}}{\Gamma(m)} = (1+d)^{-m}.$$

An alternative approach

We can find $P(X > d)$ by first finding the marginal p.d.f. of X. The joint p.d.f. of X and Y is given by

$$f(x, y) = f_Y(y) f_{X|Y}(x|y) = \frac{y^m e^{-y(1+x)}}{\Gamma(m)}, \quad y > 0, \ x > 0.$$

Therefore,

$$f_X(x) = \int_0^\infty \frac{y^m e^{-y(1+x)}}{\Gamma(m)} dy.$$

Substituting $w = y(1+x)$, we obtain

$$f_X(x) = \frac{1}{(1+x)^{m+1}\Gamma(m)} \int_0^\infty w^{m+1-1} e^{-w}\, dw = \frac{\Gamma(m+1)}{(x+1)^{m+1}\Gamma(m)}$$

$$= m(1+x)^{-m-1}, \quad x > 0.$$

Therefore,

$$P(X > d) = \int_d^\infty m(1+x)^{-m-1} dx = \left[-(1+x)^{-m}\right]_d^\infty = (1+d)^{-m}.$$

1.7.7 Conditional variance

Definition 1.42

The conditional variance of X, given the random variable Y, is defined by

$$V(X|Y) = E\left[\left\{X - E(X|Y)\right\}^2 \middle| \ Y\right]. \quad \text{...(1.7.26)}$$

Remarks

It can be seen that

$$V(X|Y) = E(X^2|Y) - [E(X|Y)]^2$$

Example 1.36

Refer to Example 1.32 on page 53. Find $V(Y|\, X = x)$.

Solution

$$E(Y^2|\, X = x) = \int_{-\infty}^\infty y^2 f_{Y|X}(y|x) dy$$

In Example 1.32, we showed that $f_{Y|X}(y|x) = \frac{1}{x}, \; x > y > 0.$ Therefore,

$$E\left(Y^2 \middle| X = x\right) = \int_0^x y^2\left(\tfrac{1}{x}\right)dy = \left[\tfrac{1}{3}y^3\left(\tfrac{1}{x}\right)\right]_0^x = \tfrac{1}{3}x^2.$$

$$E\left(Y \middle| X = x\right) = \int_0^x y\left(\tfrac{1}{x}\right) dy = \tfrac{1}{x}\left[\tfrac{1}{2}y^2\right]_0^x = \tfrac{1}{2}x, \qquad x > 0.$$

Therefore,

$$V\left(Y \middle| X = x\right) = \tfrac{1}{3}x^2 - \tfrac{1}{4}x^2 = \tfrac{1}{12}x^2.$$

1.7.8 The expected value of functions of random variables

Definition 1.43

The expected value of the random variable $Z = g(X, Y)$ is given by

$$E(Z) = \begin{cases} \sum\limits_{j=1}^{\infty}\sum\limits_{i=1}^{\infty} g(x_i, x_j)p_{X,Y}(x_i, x_j), & \text{(discrete case)}, \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f_{X,Y}(x, y)\,dxdy, & \text{(continuous case)}. \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots(1.7.27)$$

Remarks

(i) If X and Y are independent random variables and $g(X, Y) = g_1(X)g_2(Y)$, then:

$$E\{g(X, Y)\} = E\{g_1(X)g_2(Y)\} = E\{g_1(X)\}E\{g_2(Y)\}.$$

(ii) If in (i), we put $g_1(X) = X$ and $g_2(Y) = Y$, we obtain

$$E(XY) = E(X)E(Y).$$

The covariance of X and Y

Definition 1.44

The covariance of X and Y is defined by

$$\mathrm{Cov}(X, Y) = E\left\{\left(X - E(X)\right)\left(Y - E(Y)\right)\right\} = E(XY) - E(X)E(Y). \quad \ldots\ldots..(1.7.28)$$

If X and Y are independent, then $E(XY) = E(X)E(Y)$. It follows that, if X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$. However, the converse of this result is not always true (see Exercise 1(h), Question 18).

Correlation coefficient

Definition 1.45

Let X and Y be random variables with variances σ_X^2 and σ_Y^2, respectively. The correlation coefficient, ρ_{XY}, between X and Y is given by

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}. \quad \text{......(1.7.29)}$$

We have learnt that, ***if X and Y have the bivariate normal distribution, then X and Y are independent if and only if $\rho_{XY} = 0$.***

We continue with limit theorems that will be used in subsequent chapters.

Theorem 1.33

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables and let $S_n = X_1 + X_2 + \ldots + X_n$.

(a) The weak law of large numbers

If $E(X_i) = \mu < \infty$, $i = 1, 2, \ldots, n$, then

$$\lim_{n\to\infty} P\left[\left|\frac{1}{n}S_n - \mu\right| \le c\right] = 1, \ \forall\ c > 0. \quad \text{......(1.7.30)}$$

(b) The strong law of large numbers

If $E(X_i^2) < \infty$, $i = 1, 2, \ldots, n$, then

$$P\left[\lim_{n\to\infty} \frac{1}{n} S_n = \mu\right] = 1. \quad \text{......(1.7.31)}$$

(c) The central limit theorem

If $E(X_i) = \mu \in R$ and $V(X_i) = \sigma^2 < \infty$, $i = 1, 2, \ldots, n$, then

$$\lim_{n\to\infty} P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right] = \Phi(x), \quad \text{......(1.7.32)}$$

where Φ is given by Equation (1.6.10) on page 34.

The central limit theorem (CLT) implies that

$$S_n \approx N\left(n\mu,\ n\sigma^2\right) \text{ and } \tfrac{1}{n}S_n \approx N\left(\mu,\ \sigma^2/n\right), \text{ as } n \to \infty.$$

In general, for $n \geq 30$, the approximation by the normal distribution should be very good. Another application of the CLT is the [12]de Moivre-Laplace normal approximation to the binomial distribution:

$$P\left[B(n,\ p) = k\right] \approx f_Z(k), \quad \ldots\ldots(1.7.33)$$

where Z is $N\left(np,\ np(1-p)\right)$. This approximation is valid if $np > 5$ and $np(1-p) > 5$.

Example 1.37

Suppose that 1% of the tyres manufactured by a certain company do not conform to the norms (or are defective). What is the probability that among 1 000 tyres of the same type, there are exactly 10 that do not conform to the norms?

Solution

Let X be the number of tyres that do not conform to the norms among the 1 000 tyres. If we assume that the tyres are independent, then X has the binomial distribution with parameters n = 1 000 and p = 0.01. By Equation (1.7.33),

$$P\left(X = 10\right) \approx f_Z(10), \text{ where } Z \text{ is } N(10,\ 9.9).$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{9.9}} \exp\left\{-\frac{1}{2}\frac{(10-10)^2}{9.9}\right\} \approx 0.1268.$$

Remarks

(i) By using the Poisson approximation to the binomial distribution, we obtain (see Theorem 1.17)

$$P\left(X = 10\right) \approx \frac{(1\,000 \times 0.01)^{10} e^{-1\,000 \times 0.01}}{10!} = \frac{10^{10} e^{-10}}{10!} = 0.1251.$$

(ii) To calculate approximately a probability like $P(5 \leq X \leq 10)$, we would rather use the distribution function of the normal distribution. It is then recommended to make a ***continuity correction*** to improve the approximation. That is, we write that

$$P(5 \leq X \leq 10) \approx P(4.5 \leq X \leq 10.5).$$

---

[12] One of the pioneers of calculus of probabilities, Abraham de Moivre, 1667 – 1754, was born in France and died in England. The definition of independence of two events can be found in his book, *The Doctrine of Chance*, published in 1718. The formula, attributed to Stirling, appeared in a book that was published in 1730. He later used the formula to prove the normal approximation to the binomial distribution.

Exercise 1(h)

1. The joint p.d.f. of X and Y is

$$f_{X,Y}(x, y) = 6e^{-2x-3y}, \qquad x > 0, \ y > 0.$$

Find: (a) $P(X < 2, \ Y > 1)$, (b) $P(X > a)$, (c) $P(X > Y)$.

2. The joint p.d.f. of X and Y is

$$f_{X,Y}(x, y) = c(x+y), \qquad 0 < x < 1, \ 0 < y < 2,$$

where c is an unknown constant.

(a) Show that $c = \frac{1}{3}$.

(b) Find the marginal probability density functions of X and Y.

3. Let $f_{X,Y}(x, y) = \begin{cases} 6x^2 y, & 0 < x < 1, \ 0 < y < 1, \\ 0, & \text{elsewhere,} \end{cases}$

be the joint p.d.f. of X and Y.

(a) Find the marginal probability density functions of X and Y.

(b) Find: (i) $P\left(0 < X < \frac{3}{4}, \ \frac{1}{3} < Y < 1\right)$, (ii) $P\left(X > \frac{1}{2}, \ Y < \frac{1}{2}\right)$.

4. Let $f_{X,Y}(x, y) = \begin{cases} cxy, & 0 < x < 1, \ 0 \le y \le 1, \\ 0, & \text{elsewhere,} \end{cases}$

be the joint p.d.f. of X and Y.

(a) Find the value of the constant c.

(b) Find the marginal probability density functions of X and Y.

(c) Find: (i) $P\left(0 < X < \frac{1}{2}, \ \frac{1}{4} < Y < 1\right)$, (ii) $P(X < Y)$.

5. The joint p.d.f. of X and Y is

$$f_{X,Y}(x, y) = \begin{cases} c(1-y), & 0 \le x \le y \le 1, \\ 0, & \text{elsewhere,} \end{cases}$$

where c is a constant.

(a) Show that $c = 6$. (b) Find $P\left(Y < \frac{1}{2}, \ X < \frac{3}{4}\right)$. (c) Find the marginal p.d.f. of X.

6. The random vector (X, Y) has the p.d.f.

$$f_{X,Y}(x, y) = \begin{cases} c(x^2 + xy), & 0 < x < 1, \ 0 \le y \le 1, \\ 0, & \text{elsewhere,} \end{cases}$$

where c is a constant. Find:

(a) the value of the constant c, (b) $P(X > Y)$, (c) the marginal p.d.f. of X.

7. Let $f_{X|Y}(x|y) = \begin{cases} cye^{-yx}, & x > 0,\ y > 0, \\ 0, & \text{elsewhere}, \end{cases}$ and $f_Y(y) = \begin{cases} ke^{-y}, & y > 0, \\ 0, & \text{elsewhere}, \end{cases}$

denote respectively, the conditional p.d.f. of X given $Y = y$ and the marginal p.d.f. of Y. Determine:

(a) the constants c and k, (b) the joint p.d.f. of X and Y,

(c) the marginal p.d.f. of X, (d) $P(X > 1 | Y = 2)$.

8. The random vector (X, Y) has p.d.f. given by

$$f_{X,Y}(x, y) = \begin{cases} \frac{4}{9}xy, & 1 < x < 2,\ 1 < y < 2, \\ 0, & \text{elsewhere}. \end{cases}$$

(a) Find the marginal and conditional probability density functions of X and Y. Are X and Y independent?

(b) Find: (i) $P(-1 < X < 1.5 | Y > 1.5)$, (ii) $P(1.5 < Y < 2 | X < 1.4)$,

(iii) $P\left(\frac{1}{2} < X < \frac{3}{4},\ Y < \frac{3}{2}\right)$.

9. The random vector (X, Y) has p.d.f.

$$f_{X,Y}(x, y) = \begin{cases} e^{-2x}, & x \ge 0,\ 0 < y < 2, \\ 0, & \text{elsewhere}. \end{cases}$$

(a) Show that X and Y are independent. (b) Find $P\left(X < 2,\ \frac{1}{2} < Y < 1\right)$.

10. Let $f_{X,Y}(x, y) = \begin{cases} c, & -2 < x < 2,\ -1 < y < 1, \\ 0, & \text{elsewhere}, \end{cases}$

be the joint p.d.f. of X and Y.

(a) Find the value of c. (b) Determine whether X and Y are independent.

11. The random vector (X, Y) has the p.d.f.

$$f_{X,Y}(x, y) = \begin{cases} 2, & 0 \le y \le x \le 1, \\ 0, & \text{elsewhere}. \end{cases}$$

(a) Find the marginal probability density functions of X and Y.

(b) Determine whether X and Y are independent.

(c) Determine: (i) $P\left(X < \frac{1}{2} \middle| Y = \frac{1}{4}\right)$, (ii) $P\left(Y > \frac{1}{2} \middle| X = \frac{3}{4}\right)$.

12. The random vector (X, Y) has the p.d.f.

$$f_{X,Y}(x, y) = cx^2y^3, \qquad 0 < x < y < 1, \qquad \text{where } c \text{ is a constant.}$$

(a) Find the value of c. (b) Find $P\left(X > \frac{1}{4} \middle| Y = \frac{1}{2}\right)$.

13. The joint p.d.f. of X and Y is given by

$$f_{X,Y}(x, y) = \begin{cases} c, & 0 < x < y,\ 0 < y < 1, \\ 0, & \text{elsewhere,} \end{cases}$$

where c is a constant.
(a) Find the value of c.
(b) Find the marginal p.d.f. of (i) X, (ii) Y.
(c) Find the conditional p.d.f. of (i) X given $Y = y$, (ii) Y given $X = x$.

14. The random vector (X, Y) has the p.d.f.

$$f_{X,Y}(x, y) = \begin{cases} xe^{-(x+y)}, & x > 0,\ y > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the marginal probability density functions of X and Y.
(b) Find the conditional p.d.f. of Y given $X = x$. Are X and Y independent?

15. The number of accidents per year of a given bus driver has a Poisson distribution with parameter M, where M is a random variable with p.d.f.

$$f_M(m) = \lambda(\lambda m)^{k-1}e^{-\lambda m}/(k-1)!, \qquad m > 0,$$

where k is a positive integer. Find the distribution of the number of accidents of a bus driver chosen at random.

16. The conditional probability distribution of a random variable Y, given that another random variable X has taken the value x, is given by

$$P(Y = y | X = x) = \binom{10}{y} x^y (1-x)^{10-y}, \qquad y = 0, 1, \ldots, 10.$$

If the p.d.f. of X is $f_X(x) = \frac{1}{B(6,\ 4)} x^5(1-x)^3, \quad 0 < x < 1,$ find $P(Y = y)$.

17. The random variable Y has the binomial distribution with parameters $(5, X)$, where X is a random variable with p.d.f.

$$f_X(x) = \frac{1}{B(4,\ 2)} x^3(1-x), \qquad 0 < x < 1.$$

Find: (a) $E(Y)$, (b) $P(Y = y)$.

18. The following table gives the joint probability mass function of X and Y.
(a) Find: (i) Cov (X, Y), (ii) ρ_{XY}. (b) Are X and Y independent?

$f(x,y)$		y		
		-1	0	1
x	-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
	0	$\frac{1}{8}$	0	$\frac{1}{8}$
	1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

19. The moment generating function of X is given by $M_X(t) = \exp(3e^t - 3)$ and that of Y is given by $M_Y(t) = \left(\frac{1}{3}\right)^5 (2e^t + 1)^5$. If X and Y are independent, find
    (a) $P(X = 1,\ Y = 2)$, (b) $P(X = 0,\ Y = 1)$.

20. The moment generating function of X is given by
$$M_X(t) = 1/(1-2t), \quad y < \tfrac{1}{2},$$
and that of Y is given by
$$M_Y(t) = 1/(1-4t), \quad t < \tfrac{1}{4}.$$
If X and Y are independent, find $P(0 < X < 1,\ Y > 2)$.

21. The continuous random vector (X, Y) has p.d.f. given by
$$f(x,y) = e^{-y}, \quad y > x > 0.$$
(a) Find (i) $E(X^2)$, (ii) $E\left[E\left(X^2|Y\right)\right]$.
    What do you notice?
(b) Find (i) $E(Y)$, (ii) $E\left[E\left(Y|X\right)\right]$.
    What do you notice?

22. The joint p.d.f. of X and Y is given by
$$f(x,y) = \frac{1}{y}e^{-y}, \quad 0 < x < y,\ 0 < y < \infty.$$
Find: (a) $E\left(X|Y = y\right)$, (b) $E\left(X^2|Y = y\right)$, (c) $E\left[E\left(X|Y\right)\right]$,
(d) $E\left[E\left(X^2|Y\right)\right]$, (e) $V\left(X|Y = y\right)$.

23. The random variables X and Y have the joint p.d.f. given by
$$f(x,y) = cxy, \quad 0 < x < y < 1,$$
where c is a positive constant.

(a) Are X and Y independent? Give reasons for your answer.
(b) Find the constant c and the marginal and conditional probability density functions of X and Y.
(c) If $Z = E(X|Y)$, find $P\left(Z \geq \frac{1}{3}\right)$. (d) Find $P\left(X > \frac{1}{4} \middle| Y = \frac{1}{2}\right)$.

24. The random variable X has the Poisson distribution with mean M, where M is a random variable with p.d.f.

$$f_M(m) \;=\; m^{k-1}e^{-m}/(k-1)!, \qquad m > 0,$$

where k is a positive integer. By finding the moment generating function of X, find:
(a) $P(X = x)$, (b) $E(X)$. $\left(\text{Hint: } M_X(t) \;=\; E\left[E\left(e^{tX}|M\right)\right]\right)$.

25. Suppose that the random variable X is uniformly distributed over the interval (0, 1). Assume that the conditional distribution of Y given $X = x$ has a binomial distribution with parameters n and $p = x$; that is,

$$P\left(Y = y \middle| X = x\right) \;=\; \binom{n}{y} x^y (1-x)^{n-y}, \qquad y = 0, 1, \ldots, n.$$

Find: (a) $E(Y)$, (b) the probability mass function of Y.

26. Suppose X and Y have a bivariate normal distribution with $\sigma_X = 0.04$, $\sigma_Y = 0.08$, $\mu_X = 3.00$, $\mu_Y = 7.70$, and $\rho = 0$. Calculate $P(2.95 < X < 3.05,\ 7.60 < Y < 7.80)$.

27. The random variables X, Y have the joint p.d.f. given by

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right], \quad (x,\ y \in R).$$

(a) Find the distributions of X and Y.
(b) Derive the conditional p.d.f. of Y given that $X = x$.
(c) Write down $E(X)$, $E(Y)$, $E\left(X \middle| Y = y\right)$ and $V\left(Y \middle| X = x\right)$.

28. The two-dimensional random variable (X, Y) has the bivariate normal distribution with parameters $\sigma_X = 2$, $\sigma_Y = 1$, $\mu_X = 10$, and $\mu_Y = 6$.
(a) If $\rho = 0$, find $P(6 < X < 14,\ 5 < Y < 8)$. (b) If $\rho = 0.1$, find $P(5 < X < 9|Y = 2)$.

29. Let X and Y have a bivariate normal distribution with parameters $\mu_1 = 3$, $\mu_2 = 1$, $\sigma_1^2 = 16$, $\sigma_2^2 = 25$ and $\rho = \frac{3}{5}$. Determine the following probabilities.
(a) $P(3 < Y < 8)$, (b) $P(3 < Y < 8|X = 7)$, (c) $P(-3 < X < 3)$, (d) $P(-3 < X < 3|Y = -4)$.

30. (a) A random variable X has the Laplace distribution with p.d.f.

$$f_X(x) = \tfrac{1}{2} e^{-|x|}, \qquad -\infty < x < \infty.$$

Show that the characteristic function of X is given by $\phi_X(t) = 1/(1+t^2)$.

(b) Let X_1, X_2, X_3 and X_4 be independent standard normal random variables. Find:

(i) the characteristic function of $W = X_1 X_2$,

(ii) the distribution of $T = X_1 X_2 - X_3 X_4$.

31. Let X and Y have a bivariate normal distribution with parameters $\mu_1 = 5$, $\mu_2 = 10$, $\sigma_1^2 = 1$, $\sigma_2^2 = 25$ and $\rho > 0$. If $P(4 < Y < 16 | X = 5) = 0.9544$, find the value of ρ.

32. Let X and Y have a bivariate normal distribution with respective parameters $\mu_1 = 2.8$, $\mu_2 = 110$, $\sigma_1^2 = 0.16$, $\sigma_2^2 = 100$, and $\rho = 0.6$. Compute:

(a) $P(106 < Y < 124)$, (b) $P(106 < Y < 124 | X = 3.2)$.

33. We say that the continuous random variable X, whose set of possible values is the interval $[0, \infty)$, has the [13]Pareto distribution with parameter $\theta > 0$ if the density function of X is of the form

$$f_X(x) = \begin{cases} \dfrac{\theta}{(1+x)^{\theta+1}}, & x \geq 0, \\ 0, & \text{elsewhere.} \end{cases}$$

In economics, the Pareto distribution is used to describe the (unequal) distribution of wealth. Suppose that, in a given country, the wealth X of an individual (in thousands of dollars) has the Pareto distribution with parameter $\theta = 1.2$.

(a) What is the median wealth in this country?

(b) We find that about 11.65% of the population has a personal wealth of at least \$5 000, which is the average wealth in this population. What percentage of the total wealth of this country does this 11.65% of the population own?

34. The random variable X has the [14]Maxwell distribution with parameter $\alpha > 0$,

$$f_X(x) = \frac{\sqrt{2/\pi}}{\alpha^3} x^2 e^{-x^2/(2\alpha^2)}, \qquad x \geq 0.$$

[13] Vilfredo Pareto, 1848 – 1923, born in France and died in Switzerland, was an economist and sociologist. He observed that 20% of the Italian population owned 80% of the wealth of the country, which was generalized by the concept of the Pareto distribution.

[14] James Clerk Maxwell, 1831 – 1879, was born in Scotland and died in England. He was a physicist and mathematician who worked in the fields of electricity and magnetism.

Show that $E(X) = 2\alpha\sqrt{2/\pi}$ and $V(X) = \alpha^2\left[3 - (8/\pi)\right]$. This distribution is used in statistical mechanics, in particular, to describe the velocity of molecules in thermal equilibrium.

35. Let X be a continuous random variable having the probability density function

$$f_X(x) = \begin{cases} \dfrac{x}{\theta^2} e^{-x^2/(2\theta^2)}, & x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

We say that X has a [15]Rayleigh distribution with parameter $\theta > 0$.

Show that $E(X) = \theta\sqrt{\pi/2}$ and $V(X) = \theta^2\left[2 - (\pi/2)\right]$.

1.7.9 The distribution of functions of two continuous random variables

The following theorem is the two-dimensional version of Theorem 1.25.

Theorem 1.34

Suppose that (X, Y) is a two-dimensional continuous random variable with p.d.f. $f(x, y)$. Let $Z = H(X, Y)$ and $W = G(X, Y)$ be real-valued functions of X and Y. Assume that the functions H and G satisfy the following conditions:

1. The equations $z = H(x, y)$ and $w = G(x, y)$ can be solved uniquely for x and y, in terms of z and w, say $x = a(z, w)$, $y = b(z, w)$.
2. The partial derivatives $\frac{\partial x}{\partial z}, \frac{\partial x}{\partial w}, \frac{\partial y}{\partial z}$ and $\frac{\partial y}{\partial w}$ exist and are continuous.

   Then, the joint p.d.f. of Z and W, $k(z, w)$, is given by the equation

$$k(z,w) = f\left[a(z, w), b(z, w)\right]\left|J(z, w)\right|,$$

where $J = \dfrac{\partial(x, y)}{\partial(z, w)} = \begin{vmatrix} \dfrac{\partial x}{\partial z} & \dfrac{\partial x}{\partial w} \\ \dfrac{\partial y}{\partial z} & \dfrac{\partial y}{\partial w} \end{vmatrix}.$

15 John William Strutt Rayleigh, 1842 – 1919, was born and died in England. He won the Nobel Prize for physics in 1905. The distribution that bears his name is associated with the phenomenon known as Rayleigh fading in communication theory.

Theorem 1.34 can be generalized to the n-dimensional case, where $n \in \{3,\ 4,\ \ldots\}$.

Notes

1. The transformation $x = a(z,\ w),\quad y = b(z,\ w)$ is the inverse of the transformation $w = G(x,\ y),\ \ z = H(x,\ y)$.
2. J is called the ***Jacobian*** of the inverse transformation.
3. To use Theorem 1.34 to find the p.d.f. of $Z = H(X,\ Y)$, we first introduce an appropriate ***auxiliary*** random variable $W = G(X,\ Y)$. We then use the Theorem to find $k_{Z,W}(z,\ w)$, the joint p.d.f. of Z and W. The marginal p.d.f. of Z can then be found from the equation

$$f_Z(z) \ = \ \int_{-\infty}^{\infty} k_{Z,W}(z,\ w)\,dw.$$

Example 1.38

The random variables X and Y are independent with probability density functions

$$f_X(x) = e^{-x},\ \ x \geq 0, \qquad f_Y(y) = 2e^{-2y},\ \ y \geq 0.$$

Find the p.d.f. of $U = X/Y$.

Solution

Since X and Y are independent, the joint p.d.f. of X and Y is given by

$$f_{X,Y}(x,\ y) \ = \ f_X(x)\,f_Y(y) \ = \ 2e^{-x\,-\,2y}, \qquad x \geq 0,\ y \geq 0.$$

To find the p.d.f. of U, we introduce an auxiliary random variable $W = Y$. We then express X and Y in terms of U and W. This gives $x = uw,\ y = w$. The joint p.d.f. of U and W is then given by

$$h_{U,W}(u,\ w) \ = \ f_{X,Y}(uw,\ w)\left|J\right| \ = \ 2e^{-uw-2w}\left|J\right|.$$

where

$$J \ = \ \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial w} \end{vmatrix} \ = \ \begin{vmatrix} w & u \\ 0 & 1 \end{vmatrix} \ = \ w.$$

Thus,

$$k_{U,W}(u,\ w) \ = \ 2e^{-uw\,-\,2w}\left|w\right| \ = \ 2we^{-uw\,-\,2w}, \qquad u \geq 0,\ w \geq 0.$$

The p.d.f. of U is therefore given by

$$h(u) \ = \ \int_0^{\infty} 2we^{-w(u\,+\,2)}dw.$$

7. X_1 and X_2 are independent χ^2 random variables, with r_1 and r_2 degrees of freedom, respectively. Find the p.d.f. of $Y = X_1 + X_2$.

8. Let X and Y be independent random variables with probability density functions:

$$f_X(x) = x^{a-1}(1-x)^{b-1}/B(a,\ b), \qquad 0 < x < 1,\ a > 0,\ b > 0,$$
$$f_Y(y) = y^{d-1}e^{-y}/\Gamma(d), \qquad y > 0, \quad d > 0.$$

Show that $U = XY$ and $V = (1-X)Y$ are independent if $a + b = d$, and find their distributions.

9. If X and Y have the joint p.d.f.

$$f(x, y) = \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{x^2 - 2\rho xy + y^2\right\}\right],$$

where $|\rho| < 1$, $-\infty < x < \infty$ and $-\infty < y < \infty$, find the p.d.f. of $U = X/Y$.

10. If X and Y are independent standard normal variables, find the p.d.f. of $U = X/Y$.

11. The random vector (X, Y) has the p.d.f. given by

$$f(x, y) = c(x+y), \qquad 0 < x < 2,\ 0 < y < 2.$$

(a) Show that $c = \frac{1}{8}$. (b) Find the marginal p.d.f. of Y.

(b) Let $U = X + Y$ and $V = Y$. Obtain $h(u,\ v)$, the joint p.d.f. of U and V, stating the values of u and v for which $h(u,\ v)$ is non-zero. Use the joint p.d.f. of U and V to obtain the marginal p.d.f. of V, and confirm that it is indeed the same as the p.d.f. of Y.

12. Let X and Y be independent random variables with probability density functions

$$f_X(x) = 12x^2(1-x), \quad 0 < x < 1, \qquad f_Y(y) = y^4 e^{-y}/4!, \quad y > 0.$$

Find the probability density functions of $U = XY$ and $V = (1-X)Y$.

13. (a) If W has the [16]Laplace distribution with p.d.f.

$$f_W(w) = \tfrac{1}{2}e^{-|w|}, \qquad -\infty < w < \infty,$$

find the characteristic function of W.

(b) X and Y are independent identically distributed exponential random variables with mean 2. Find the characteristic function of $U = \frac{1}{2}(X - Y)$ and deduce the distribution of U.

---

[16] See page 46.

14. Let X_1, X_2 and X_3 have the joint p.d.f.

$$f(x_1, x_2, x_3) = \begin{cases} 48x_1x_2x_3, & 0 < x_1 < x_2 < x_3 < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

If $Y_1 = X_1/X_2$, $Y_2 = X_2/X_3$ and $Y_3 = X_3$, show that Y_1, Y_2 and Y_3 are independent and find their probability density functions.

15. Let X_1, X_2 and X_3 be *i.i.d.* with common p.d.f. $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$, zero elsewhere. Let

$$Y_1 = \frac{X_1}{X_1 + X_2}, \quad Y_2 = \frac{X_1 + X_2}{X_1 + X_2 + X_3} \text{ and } Y_3 = X_1 + X_2 + X_3.$$

Find the probability density functions of Y_1, Y_2 and Y_3 and show that Y_1, Y_2 and Y_3 are independent.

16. Let X_1, X_2, X_3 and X_4 have the joint p.d.f. $f(x_1, x_2, x_3, x_4) = 24$, $0 < x_1 < x_2 < x_3 < x_4 < 1$, 0, elsewhere. Find the joint p.d.f. of $Y_1 = X_1/X_2$, $Y_2 = X_2/X_3$, $Y_3 = X_3/X_4$ and $Y_4 = X_4$ and show that they are independent. Find the probability density function of Y_1, Y_2 and Y_3.

References

David, F. N. (1962). Games, Gods and Gambling. *Hafner, New York*.

Ofosu, J. B. and Hesse, C. A. (2011). Introduction to probability and probability distributions. *EPP Books Services, Accra*.

Todhunter, I. (1931). A History of Mathematical Theory of Probability. *G. E. Stechert, New York*.

the value of $Q(t)$ decreases by one. Thus, as t varies, the value of $Q(t)$ changes. This family of random variables, $\{Q(t):\ t\geq 0\}$, is another example of a stochastic process.

Definition 2.1

A family (or collection) of random variables that are indexed by a parameter, such as time, is called a stochastic process.

Example 2.4

Let $Y(t)$ be the volume of water in a dam at time t. Then $\{Y(t):\ t\geq 0\}$, is a stochastic process.

Example 2.5

Let $X(t)$ be the average volume of water in a dam in year t. If the dam was constructed in say 1940, then $\{X(t):\ t\in T\}$ is a stochastic process, where $T=\{1940,\ 1941,\ ...\}$.

2.2 Specification of stochastic processes

The main elements distinguishing between stochastic processes are the nature of its state space, its index set T, and the dependence relations among the random variables $X(t)$.

(a) States

The values which $X(t)$ can take, are called the ***states*** of $X(t)$. In Example 2.1, the states of the stochastic process $X(n)$ are 1, 2, 3, 4, 5, and 6, and in Example 2.3, the states of the stochastic process $\{Q(t):\ t\geq 0\}$ are $0,\ 1,\ 2,\ \cdot\ \cdot\ \cdot$. Changes in the values of the stochastic process $X(t)$ are called ***transitions*** between the states of $\{X(t)\}$. If $X(t)=i$, then the process is said to be in state i at time t.

(b) The state space, *S*

Definition 2.2

The set of possible values of $X(t)$ is called the state space of $\{X(t)\}$. It is denoted by S.

In Example 2.1, $S=\{1,\ 2,\ \cdot\ \cdot\ \cdot,\ 6\}$ and in Example 2.3, $S=\{0,\ 1,\ 2,\ \cdot\ \cdot\ \cdot\}$. If the state space of $\{X(t)\}$ is discrete, then $\{X(t)\}$ is called a ***discrete state stochastic process***. If the state space of $\{X(t)\}$ is continuous, then $\{X(t)\}$ is called a ***continuous state stochastic process***. It can be seen that the stochastic processes given in Examples 2.1 to 2.3 are discrete state stochastic processes, while the stochastic processes given in Examples 2.4 and 2.5 are continuous state stochastic processes.

(c) The index set, T

If T is discrete, then $\{X(t): t \in T\}$ is called a ***discrete time stochastic process*** and in this case, it is customary to use n instead of t, and write $X(n)$ or X_n instead of $X(t)$. If T is continuous, then $\{X(t)\}$ is called a ***continuous time stochastic process***. Examples 2.1 and 2.2 are discrete time stochastic processes while Examples 2.3 and 2.4 are continuous time stochastic processes. Example 2.5 is a continuous state and a discrete time stochastic process.

Every stochastic process can be specified in terms of its state space S, and its index set T. The following four examples are given so that the reader gets a better understanding of the concept of a stochastic process, its state space and its index set.

Example 2.6

John and Ann play a coin tossing game. John agrees to pay Ann $1.00 whenever the coin and S_n be the amount earned by Ann in n tosses of the coin.

(a) Determine the state space and the index set of the stochastic process $\{S_n\}$.

(b) If the coin is fair, show that $E(S_n)=0$ and find $V(S_n)$.

Solution

(a) Here, $n=\{1, 2, 3, \ldots\}$ and $S_n=\{0, \pm 1, \pm 2, \ldots, \pm n\}$. Thus, $\{S_n\}$ is a discrete state stochastic process indexed by a discrete set.

(b) Let $X(i)$ denote the amount, in dollars, earned by Ann in the i^{th} toss of the coin. Then,

$$X(i) = \begin{cases} 1 & \text{with probability } \frac{1}{2}, \\ -1 & \text{with probability } \frac{1}{2}. \end{cases}$$

Thus, $E\left(X(i)\right) = 1\times\frac{1}{2} + (-1)\times\frac{1}{2} = \frac{1}{2} - \frac{1}{2} = 0$, $E\left[\left\{X(i)\right\}^2\right] = 1^2\times\frac{1}{2}+\left(-1\right)^2\times\frac{1}{2} = 1$,

and so $V\left(X(i)\right)=1$. It follows that:

$$E\left(S_n\right) = \sum_{i=1}^{n} E\left(X(i)\right) = 0, \quad \text{and} \quad V\left(S_n\right) = \sum_{i=1}^{n} V\left(X(i)\right), \text{ by independence}$$
$$= n.$$

Example 2.7

Five green balls and 3 white balls are placed in two boxes A and B so that each box contains 4 balls. At each stage, a ball is drawn at random from each box and the two balls are interchanged. Let $X(n)$ denote the number of white balls in box A after the n^{th} draw. The state space and the index set of the stochastic process $\left\{X(n)\right\}$ are, respectively, $S = \left\{0, 1, 2, 3\right\}$ and $T = \left\{1, 2, 3, ...\right\}$.

If Y_n is the total number of green balls in box A after the n^{th} draw, then the stochastic process $\left\{Y_n\right\}$ has the state space $S = \left\{1, 2, 3, 4\right\}$ and index set $T=\left\{1, 2, 3, ...\right\}$. Both $\{X(n)\}$ and $\{Y_n\}$ are discrete state and discrete time stochastic processes.

Example 2.8

Consider the number of telephone calls received at a telephone exchange at time t. This can be represented by the stochastic process $\left\{X(t):\, t\geq 0\right\}$, where $X(t)$ is the number of telephone calls received at time t. This is a discrete state and a continuous time stochastic process.

The index set of each of the stochastic processes we have considered so far, is one dimensional. In the following example, we consider a stochastic process whose index set is two dimensional.

Example 2.9

Consider waves in oceans. Let $X(t)$ denote the height of the wave at the location t. We may regard the latitude and longitude coordinates of the wave as the value of t. We then have a stochastic process whose index set is not one dimensional.

2.3 Properties of stochastic processes

We now describe some of the classical types of stochastic processes, characterized by different dependence relationship among $X(t)$.

(a) Process with independent increments

Given a stochastic process $\{X(t)\}$, if the random variables $X(t_2)-X(t_1)$, $X(t_3)-X(t_2)$, , $X(t_n) - X(t_{n-1})$, are independent for all choices of t_1, t_2, ..., t_n, satisfying $t_1 < t_2 < ... < t_n$, then we say that $\{X(t)\}$ is a stochastic process with ***independent increments***.

(b) Process with the Markovian property

Consider the following example.

Example 2.10

Suppose a coin is tossed ten times and X_n denotes the total number of heads which appear up to the n^{th} toss. This gives the stochastic process $\{X_n : n=1, 2, \cdot\cdot\cdot\}$ with state space $S = \{0, 1, 2, \cdot\cdot\cdot, 10\}$ and index set $T=\{1, 2, 3, \cdot\cdot\cdot, 10\}$. It can be seen that:

$$P\left(X_3=2 \middle| X_2=1, X_1=1\right) = P\left(X_3=2 \middle| X_2=1\right).$$

Furthermore,

$$P\left(X_3=2 \middle| X_2=1, X_1=0\right) = P\left(X_3=2 \middle| X_2=1\right),$$

and $$P\left(X_3=2 \middle| X_2=2, X_1=1\right) = P\left(X_3=2 \middle| X_2=2\right).$$

Notice that these conditional probabilities depend only on the values of X_2 and are independent of the values of X_1. It can also be seen that:

$$P\left(X_4=3 \middle| X_3=3, X_2=2, X_1=1\right) = P\left(X_4=3 \middle| X_3=3\right),$$
$$P\left(X_4=3 \middle| X_3=2, X_2=2, X_1=1\right) = P\left(X_4=3 \middle| X_3=2\right),$$
$$P\left(X_4=3 \middle| X_3=2, X_2=1, X_1=0\right) = P\left(X_4=3 \middle| X_3=2\right),$$

and $$P\left(X_4=4 \middle| X_3=3, X_2=2, X_1=1\right) = P\left(X_4=4 \middle| X_3=3\right).$$

Notice that the conditional probabilities depend only on the values of X_3 and not on the values of X_1 and X_2. In general,

$$P\left(X_{n+1}=i+1 \middle| X_n=i,\ X_{n-1}=k_{n-1},\ \cdots, X_1=k_1\right) = P\left(X_{n+1}=i+1 \middle| X_n=i\right),$$

and $$P\left(X_{n+1}=i \middle| X_n=i,\ X_{n-1}=k_{n-1},\ \cdots, X_1=k_1\right) = P\left(X_{n+1}=i \middle| X_n=i\right).$$

This conditional probability depends only on the value of X_n and not on the values of $X_1, X_2, \cdots, X_{n-1}$. Thus, when X_n is known, X_{n+1} does not depend on $X_1, X_2, \cdots, X_{n-1}$. When this happens, the stochastic process $\{X_n\}$ is said to have the ***Markovian property***. It can be seen that the stochastic process $\{X_n\}$ has the Markovian property if, given the value of X_t, the values of X_s for $s>t$, are not affected by the values of X_u for $u<t$. A stochastic process which has the Markovian property is called a ***Markov process***. We now give a formal definition of a Markov process.

Definition 2.3

A stochastic process $\{X_n\}$ is a Markov process if

$$P\left(X_{n+1}=k_{n+1} \middle| X_n=k_n,\ X_{n-1}=k_{n-1},\ \cdots, X_1=k_1\right) = P\left(X_{n+1}=k_{n+1} \middle| X_n=k_n\right).$$

If a stochastic process $\{X_n\}$ has the Markovian property, then given the present state X_n, the past states $X_1, X_2, \cdots, X_{n-1}$ are not needed to predict the future state X_{n+1}.

(c) Process with stationary increments

Definition 2.4

A stochastic process $\{X(t);\ t\in T\}$ is said to have stationary increments if the distribution of the increment, $X(t_1+h)-X(t_1)$, depends only on the length h of the interval and not on the time t.

For a process with stationary increments, the distribution of $X(t_1+h)-X(t_1)$ is the same as the distribution of $X(t_2+h)-X(t_2)$, no matter the values of t_1, t_2 and h. If $\{X_t\}$ is a stochastic process with stationary increments, then the distribution of $X(t)$ is the same for each t. This also means that the particular times at which we examine the process is irrelevant.

2.4 Some common stochastic processes

We now briefly list some special stochastic processes. Each of these processes has its properties and applications.

(a) The [1]Bernoulli process

The stochastic process $\{X_n : n=1, 2, ...\}$ is called a ***Bernoulli process*** with success probability p if

(i) $X_1, X_2, \ldots$ are independent;

(ii) $P(X_n = 1) = p, \quad P(X_n = 0) = 1-p = q,$ for all n.

The Bernoulli process $\{X(n)\}$ has state space $S = \{0, 1\}$ and index set $T = \{1, 2, 3, ...\}$.

Example 2.11

Due to some random causes, the output from a certain manufacturing process occasionally turns out some defective items. Suppose the proportion of defective items is p. Let D_n denote the number of defectives found in n manufactured items. Then, $\{D_n : n=0, 1, 2, ...\}$ is a Bernoulli process with parameter p.

(b) The binomial process

Let $\{X_n : n=1, 2, ...\}$ be a Bernoulli process with success probability p and let

$$S_n = \begin{cases} 0, & n=0, \\ X_1 + X_2 + \cdots + X_n, & n = 1, 2, \ldots . \end{cases}$$

Then, S_n is the number of successes in the first n Bernoulli trials. Thus,

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n.$$

The stochastic process $\{S_n : n=0, 1, 2, ...\}$ is called the ***binomial process***. The state space of the binomial process $\{S_n\}$ is $S = \{0, 1, 2, \ldots, n\}$ while the index set is T

[1] See page 21.

(c) The [2]Poisson process

A Poisson process of intensity $\lambda > 0$ is an integer-valued stochastic process $\{N(t): t \geq 0\}$ with the following properties:

(i) $N(0) = 0$,

(ii) For $s > t,\ t > 0,$ the random variable $N(t+s) - N(t)$ has the Poisson distribution with parameter λs.

(iii) $N(t)$ has independent increments.

The state space of the Poisson process $\{N(t): t \geq 0\}$, is $S = \{0, 1, 2, ...\}$. This is discrete and the index set T is continuous. The Poisson process arises in many applications, especially as a model for the arrival at a service point such as the arrival of customers to a bank, the arrival of calls at a telephone exchange, etc.

(d) The Gaussian process

A stochastic process $\{X(t): t \geq 0\}$ is said to be Gaussian if the n-dimensional random variable $\{X(t_1), \cdot\cdot\cdot, X(t_n)\}$ has the multivariate normal distribution for all $n \geq 1$ and all $t_1, \cdot\cdot\cdot, t_n \in [0, \infty)$. This process has a continuous state space and a continuous index set. In electrical engineering, Gaussian processes appear as models of noise voltages in resistors and models of receiver noises in communication systems.

(e) The [3]Wiener process

The Wiener process is a stochastic process $\{W(t): t \geq 0\}$ with the following properties:

(i) $W(0) = 0.$

(ii) $\{W(t)\}$ has stationary and independent increments.

(iii) For all $t > 0,\ W(t)$ is $N(0,\ c^2 t),$ where c is a constant.

The process has a continuous state space and a continuous index set. It is sometimes called the ***Brownian motion process*** and has applications in quantum mechanics, diffusion phenomena, economics, etc.

---

[2] See page 22.

[3] Norbert Wiener, 1894 1964, was born in the United States and died in Sweden. He obtained his PhD degree in philosophy from Harvard University at the age of 18. His research subject was mathematical logic. After a stay in Europe to study mathematics, he started working at the Massachusetts Institute of Technology, where he did some research on Brownian motion. He contributed, in particular, to communication theory and to control.

.

Bibliographic Notes

The topics introduced in this chapter are nearly all dealt with in more detail in later chapters and references will be given there. Bartlett (1953) has given a brief and elementary introduction to a wide range of applications of the theory of stochastic processes. Neyman (1960) has discussed, with examples, the general role of stochastic processes in science. Cramér (1964) has made a broad survey of model building with stochastic processes. A very useful collection of exercises on stochastic processes has been published by Takács (1960). For an introduction to the general theory of stochastic process, see the books by Rosenblatt (1962) and Pitt (1963).

Exercise 2(a)

1. Consider the experiment in which we record $M(t)$, the number of active calls at a telephone switch at time t, for each second over an interval of 15 minutes. Determine the state space and the index set of the stochastic process $\{M(t): t \geq 0\}$.

2. Six green balls and 4 white balls are placed in two boxes A and B such that each box contains 5 balls. At each stage, a ball is drawn at random from each box and the two balls are interchanged.
   (a) Let X_n denote the number of white balls in box A after the n^{th} draw. Find the state space and the index set of the stochastic process $\{X_n\}$.
   (b) If Y_n is the total number of green balls in box A after the n^{th} draw, find the state space and the index set for the stochastic process $\{Y_n\}$.

3. A box contains 3 black and 7 white balls. At each trial, a ball is drawn randomly from the box. If it is white, it is put back into the box and if it is black, it is kept outside the box. Let X_n denote the number of black balls taken out of the box after the n^{th} trial. Find the state space and the index set of the stochastic process $\{X_n\}$.

4. Suppose a coin is tossed 15 times. Let X_n

   appear up to the n^{th} toss. Find the state space and the index set of the stochastic process $\{X_n\}$.

5. Consider an acceptance sampling plan. Let D_n denote the number of defective items after inspecting n items. Find the state space and the index set of the stochastic process $\{D_n\}$.

6. M coins are placed in a row on a table. At each stage, a coin is selected at random and turned over. Let X_n M coins after the n^{th} trial. Find the state space and the index set of the stochastic process $\{X_n\}$.

7. Two white balls and 5 green balls are placed in two boxes A and B so that Box A contains 4 balls and Box B contains 3 balls. At each stage, one ball is selected at random from each box and the two balls are interchanged. Let Y_n denote the total number of white balls in Box A at the n^{th} trial. Find the state space and the index set of the stochastic process $\{Y_n\}$.

8 John and Alice start a game with $3 each as their capital. At the end of every game, the loser pays $1 to the winner. With every game played, the probability of John winning is 0.6 and the probability of Alice winning is 0.4. They quit playing when one of them losses all his/her capital. Let C_n be the balance of John after n games. Find the state space and the index set of the stochastic process $\{C_n\}$.

References

Bartlett, M. S. (1953). Stochastic Processes or the Statistics of Change, *Appl. Statistics*, **2**, 44 64.

Cramér, R. (1964). Model building with the aid of stochastic processes. *Technometrics*, **6**, 133 160.

Neyman, J. (1960). Indetermination in science and new demands on statisticians, *J. Amer. Statist. Ass.* **55**, 625 639.

Pitt, N. R. (1963). Integration, measure and probability. *Edinburgh: Oliver and Boyd*.

Rosenblatt, M. (1962). Random processes. *Oxford University Press, New York.*

Takács, L. (1960). Stochastic process, problems and solutions. *Methuen. London*.