MATHEMATICAL STATISTICS
Lecture Note

Akram Al-Sabbagh
Al-Nahrain University

CONTENTS

1 Introduction to Statistics **1**

1.1 Basic Probability 1
1.1.1 Sample Space 1
1.1.2 Sample Event 1
1.1.3 Probability Axioms 2
1.1.4 Conditional Probability 3
1.1.5 Independence 3
1.2 Random Variables 4

2 Distribution of Random Variables **7**

2.1 The Probability Density Function (PDF) 7
2.2 Cumulative Distribution Function (CDF) 14
2.3 Mathematical Expectation 20
2.3.1 Some Special Mathematical Expectations 22
2.3.2 Tchebyshev's Inequality 29
2.4 Transformation of Variables (cdf technique) 31

3 Some Special Mathematical Distributions **35**

3.1 Discrete Distributions 35
3.1.1 Binomial Distribution 35
3.1.2 The Geometric Distribution 39

3.1.3 Negative Binomial Distribution 41
3.1.4 The Hypergeometric Probability Distribution 42
3.1.5 The Poisson Distribution 44
3.2 Continuous Distribution 46
3.2.1 The Uniform Distribution 46
3.2.2 Gamma Distribution 48
3.2.3 The Exponential Distribution 51
3.2.4 Chi-Square Distribution 53
3.2.5 The Beta Distribution 54
3.2.6 The Normal Distribution 56

4 Multivariate Probability Distributions 61

4.1 Introduction 61
4.2 Bivariate and Multivariate Probability Distributions 61
4.2.1 Joint Probability Density Function 62
4.2.2 Joint Cumulative Distribution Function 66
4.2.3 Marginal Distribution 68
4.2.4 Conditional Distribution 73
4.2.5 Independent Random Variables 76
4.3 Expected Values, Covariance, and Correlation 77
4.3.1 The Covariance and Correlation Coefficient 79

5 Function of Random Variables 83

5.1 Introduction 83
5.2 The Joint pdf of a Function of Random Variables 84
5.2.1 The Method of Transformations 84
5.3 Order Statistics 89
5.3.1 The distribution of Order Statistics 90
5.3.2 Joint Distribution of the Order Statistics 92

6 Sampling Distributions 95

6.1 Introduction 95
6.2 Sampling Distributions Related to the Normal Distribution 96
6.3 The Central Limit Theorem 99

7 Point Estimation 101

7.1 Introduction 101
7.2 General Concepts and Criteria 101
7.2.1 The Bias and Mean Square Error of Point Estimators 102
7.2.2 Some Common Unbiased Point Estimators 103
7.3 Methods of Point Estimation 105
7.3.1 The method of Moment 105
7.3.2 Maximum Likelihood Method 106

CHAPTER 1

INTRODUCTION TO STATISTICS

1.1 Basic Probability

This chapter is a reminder of some basics in probability and statistics, it contains some definitions and examples that we will be using in the rest of this notes. Probability or chance can be measured on a scale which runs from zero to one, where zero represents impossibility and one represents certainty.

1.1.1 Sample Space

A sample space, Ω, is the set of all possible outcomes of an experiment. The sample space can be classified into two main categories: discrete, where the space contains a finite or countable number of distinct point, and continuous when the space contains an uncountable distinct sample points. An event E is defined to be a subset of the sample space, $E \subseteq \Omega$.

1.1.2 Sample Event

An event is any collection (subset) of outcomes contained in the sample space Ω . An event is said to be simple if it consists of exactly one outcome and compound if it consists of more than one outcome.

EXAMPLE 1.1

A manufacturer has five seemingly identical computer terminals available for shipping. Unknown to her, two of the five are defective. A particular order calls for two of the terminals and is filled by randomly selecting two of the five that are available.

a List the sample space for this experiment.

Please enter \offprintinfo{(Title, Edition)}{(Author)}
at the beginning of your document.

b Let A denote the event that the order is filled with two non defective terminals. List the sample points in A.

c List the possible outcome for event B where both terminals are defective.

d Let C represent the case where at least one of the terminals is defective.

Solution.

a Let the two defective terminals be labelled D_1 and D_2 and let the three good terminals be labelled G_1, G_2, and G_3. Any single sample point will consist of a list of the two terminals selected for shipment. The simple events may be denoted by

$$\begin{array}{llll} E_1 = \{D_1, D_2\}, & E_5 = \{D_2, G_1\}, & E_8 = \{G_1, G_2\}, & E_{10} = \{G_2, G_3\}. \\ E_2 = \{D_1, G_1\}, & E_6 = \{D_2, G_2\}, & E_9 = \{G_1, G_3\}, & \\ E_3 = \{D_1, G_2\}, & E_7 = \{D_2, G_3\}, & & \\ E_4 = \{D_1, G_3\}, & & & \end{array}$$

Thus, The sample space Ω contains 10 sample points $\Omega = \{E_1, E_2, \ldots, E_{10}\}$.

b Event $A = \{E_8, E_9, E_{10}\}$.

c $B = \{E_1\}$.

d $C = \{E_1, E_2, E_3, E_4, E_5, E_6, E_7\}$

◀

1.1.3 Probability Axioms

Suppose Ω is a sample space associated with an experiment. To every event A in Ω (A is a subset of Ω), we assign a number, $P(A)$, called the probability of A, so that the following axioms hold:

1. $P(A) \geq 0$.
2. $P(\Omega) = 1$.
3. If $A_1, A_2, A_3, \ldots$ form a sequence of pairwise mutually exclusive events in Ω ($A_i \cap A_j = \emptyset$ for $i \neq j$), then:

$$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = \sum_{i=1}^{\infty} P(A_i).$$

Other Consequences:

(i) $P(\bar{A}) = 1 - P(A)$, therefore $P(\phi) = 0$.

(ii) For any two events A_1 and A_2 we have the addition rule:

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

.

EXAMPLE 1.2

Following example (1.1), evaluate:

a Assign probabilities to the simple events in such a way that the information about the experiment is used.

b Find the probability of event A, B, C.

c Find the probability of $A \cap B$, $A \cup C$ and $B \cap C$.

Solution. **a** Because the terminals are selected at random, any pair of terminals is as likely to be selected as any other pair. Thus, $P(E_i) = 1/10$, for $i = 1, 2, \ldots, 10$, is a reasonable assignment of probabilities.

b Since $A = E_8 \cup E_9 \cup E_10$, then $P(A) = P(E_8) + P(E_9) + P(E_{10}) = 3/10$.
Also, $P(B) = 1/10$ and $P(C) = 7/10$.

c Since $A \cap B = \emptyset$, then $P(A \cap B) = 0$.
$A \cup C = \Omega$, then $P(A \cup C) = 1$.
$B \cap C = E_1$, then $P(B \cap C) = 1/10$.

◀

1.1.4 Conditional Probability

Suppose $P(A_2) \neq 0$. The conditional probability of the event A_1 given that the probability of the event A_2 is known, is defined as:

$$P(A_1|A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}.$$

The conditional probability is undefined if $P(A_2) = 0$. The conditional probability formula above yields the multiplication rule:

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1)P(A_2|A_1) \\ &= P(A_2)P(A_1|A_2) \end{aligned} \tag{1.1}$$

1.1.5 Independence

Suppose that events A_1 and A_2 are in sample space Ω, A_1 and A_2 are said to be independent if

$$P(A_1 \cap A_2) = P(A_1)P(A_2).$$

In the case of conditional probability, this implies to $P(A_1|A_2) = P(A_1)$ and $P(A_2|A_1) = P(A_2)$. That means that the knowledge of the occurrence of one of the events does not affect the likelihood of occurrence of the other. For more general case, $A_1, A_2, \ldots$ are pairwise independent if $P(A_i \cap A_j) = P(A_i)P(A_j)$, for all $i \neq j$. They are mutually independent if for all subsets $P(\cap_j A_j) = \prod_j P(A_j)$.

EXAMPLE 1.3

Back again to example (1.1), evaluate the probability of the event A given B and B given C.

Solution.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = 0.$$

$$P(B|C) = \frac{P(B \cap C)}{P(C)} = \frac{1/10}{7/10} = 1/7.$$

◀

1.2 Random Variables

A ***Random variable*** X is a real-valued function for which the domain is a sample space. Given a random experiment with sample space Ω, then $X : \Omega \to \mathbb{R}$. The space of the r.v X is the set of real numbers $\mathcal{A} = \{x : x = X(\omega), \omega \in \Omega\}$. Furthermore, for any event $A \subset \mathcal{A}$, then there is an event $\Psi \subset \Omega$, such that $P(A) = \Pr\{X \in A\} = P(\Psi)$, where $\Psi = \{\omega : \omega \in \Omega, X(\omega) \in A\}$ and $A = \{x : x = X(\omega), \omega \in \Psi\}$, knowing that $P(A)$ satisfy the probability axiom 1.1.3. ***Note:*** A r.v X is called discrete if it defined on a discrete sample space (countable or finite), and it is called a continuous r.v otherwise.

EXAMPLE 1.4

When a student attempts to connect to a university computer system, either there is a failure (F), or there is a success (S). With $\omega = S, F$, define an rv X by $X(S) = 1$, $X(F) = 0$. The rv X indicates whether (1) or not (0) the student can connect.

EXAMPLE 1.5

Starting at a fixed time, each car entering an intersection is observed to see whether it turns left (L), right (R), or goes straight ahead (A). The experiment terminates as soon as a car is observed to turn left. Let $X =$ the number of cars observed.

Therefore X can takes values from zero (if the first car turned the left) to infinity (if non of the cars turned left). then $x = 0, 1, 2, 3, \ldots$.

EXAMPLE 1.6

The number of pumps in use at both a six-pump station and a four-pump station will be determined. Then the possible values for each of the following random variables are:

1. $T =$ the total number of pumps in use. Then $T = 0, 1, 2, \ldots, 10$.
2. $X =$ the difference between the numbers in use at stations 1 and 2. Then $X = 0, 1, \ldots, 6$.
3. $Z =$ the number of stations having exactly two pumps in use. Then $Z = 0, 1, 2$.

EXAMPLE 1.7

Toss a coin twice, the the sample space is: $\Omega = \{HH, HT, TH, TT\}$, suppose a r.v X represent the number of heads. Then:

$$X(\omega) = \begin{cases} 0, & \text{if } \omega = TT \\ 1, & \text{if } \omega = TH, HT \\ 2, & \text{if } \omega = HH \end{cases} \tag{1.2}$$

Therefore, the space of X is $\mathcal{A} = \{x : x = 0, 1, 2\}$, and the probability of $x = 0, 1, 2$: $\Pr\{X = 0\} = 1/4$, $\Pr\{X = 1\} = 1/2$ and $\Pr\{X = 2\} = 1/4$.

Assume the event $A = \{x : x = 0, 1\} \subset \mathcal{A}$, then $P(A) = \Pr(X \in A) = \Pr(X = 0, 1) = \Pr(X = 0) + \Pr(X = 1) = 3/4$.

EXAMPLE 1.8

If in the study of the ecology of a lake, we make depth measurements at randomly chosen locations, then $X =$ the depth at such a location is a continuous rv. Here A is the minimum depth in the region being sampled, and B is the maximum depth.

EXAMPLE 1.9

If a chemical compound is randomly selected and its pH Z is determined, then Z is a continuous rv because any pH value between 0 and 14 is possible. If more is known about the compound selected for analysis, then the set of possible values might be a subinterval of $[0, 14]$, such as $5.5 < z < 6.5$, but Z would still be continuous.

EXAMPLE 1.10

Let X represent the amount of time a randomly selected customer spends waiting for a haircut before his/her haircut commences. Your first thought might be that X is a continuous random variable, since a measurement is required to determine its value. However, there are customers lucky enough to have no wait whatsoever before climbing into the barber's chair. So it must be the case that $P(X = 0) > 0$. Conditional on no chairs being empty, though, the waiting time will be continuous since X could then assume any value between some minimum possible time A and a maximum possible time B.

EXAMPLE 1.11

Let $\mathcal{A} = \{x : 0 < x < 2\}$ be the sample space of a r.v X. For each event $A \subset \mathcal{A}$, we define the probability set function $P(A)$ as

$$P(A) = \begin{cases} \int\limits_{x \in A} \frac{3}{8} x^2 dx, & x \in \mathcal{A} \\ 0, & e.w \end{cases} \tag{1.3}$$

If $A_1 = \{x : 0 < x < 1/2\}$ and $A_2 = \{x : 1/2 < x < 1\}$. Find the $P(A_1), P(A_1^c), P(A_2), P(A_2^c), P(A_1 \cap A_2), P(A_1 \cup A_2)$

Solution.

$$P(A_1) = \int\limits_{x \in A_1} \frac{3}{8} x^2 dx = \frac{3}{8} \int_0^{1/2} x^2 dx = \frac{1}{64}.$$

$$P(A_1^c) = 1 - P(A_1) = 1 - \frac{1}{64} = \frac{63}{64}.$$

$$P(A_2) = \int\limits_{x \in A_2} \frac{3}{8} x^2 dx = \frac{3}{8} \int_{1/2}^{1} x^2 dx = \frac{7}{64}.$$

$$P(A_2^c) = 1 - P(A_2) = 1 - \frac{7}{64} = \frac{57}{64}.$$

Since $A_1 \cap A_2 = \emptyset$, Then $P(A_1 \cap A_2) = 0$, and then

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) = \frac{8}{64} = \frac{1}{8}.$$

◀

EXAMPLE 1.12

Let $\mathcal{A} = \{x : x = 1, 2, \dots\}$ be the sample space of a r.v X. For each event $A \subset \mathcal{A}$, we define the probability set function $P(A)$ as

$$P(A) = \Pr(X \in A) = \sum_{x \in A} \left(\frac{1}{2}\right)^x, \quad x \in \mathcal{A}$$

If $A = \{x : x = 1, 2\}, B = \{x : x = 2, 3\}, C = \{x : x = 1, 3, 5, \dots\}$. Find $P(A),\ P(B),\ P(C),\ P(A^c),\ P(B^c),\ P(C^c),\ P(A \cap B)$, and $P(A \cup B)$.

Solution.

$$P(A) = \sum_{x \in A} \left(\frac{1}{2}\right)^x = \sum_{x=1}^{2} \left(\frac{1}{2}\right)^x = \frac{1}{2} + \left(\frac{1}{2}\right)^2 = \frac{3}{4} \Rightarrow P(A^c) = \frac{1}{4}$$

$$P(B) = \sum_{x \in B} \left(\frac{1}{2}\right)^x = \sum_{x=2}^{3} \left(\frac{1}{2}\right)^x = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 = \frac{3}{8} \Rightarrow P(B^c) = \frac{5}{8}$$

$$P(C) = \sum_{x \in C} \left(\frac{1}{2}\right)^x = \sum_{x=1, step2}^{\infty} \left(\frac{1}{2}\right)^x = \frac{1}{2} + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^5 + \cdots = \frac{1/2}{1 - (1/2)^2} = \frac{2}{3} \Rightarrow P(C^c) = \frac{1}{3}$$

$$A \cap B = \{x : x = 2\} \Rightarrow P(A \cap B) = \sum_{x \in A \cap B} \left(\frac{1}{2}\right)^x = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) == \frac{3}{4} + \frac{3}{8} - \frac{1}{4} = \frac{7}{8}$$

◀

CHAPTER 2

DISTRIBUTION OF RANDOM VARIABLES

2.1 The Probability Density Function (PDF)

Notationally, we will use an upper case letter, such as X or Y, to denote a random variable and a lower case letter, such as x or y, to denote a particular value that a random variable may assume. For example, let X denote any one of the six possible values that could be observed on the upper face when a die is tossed. After the die is tossed, the number actually observed will be denoted by the symbol x. Note that X is a random variable, but the specific observed value, x, is not random.

It is now meaningful to talk about the probability that X takes on the value x, denoted by $\Pr(X = x)$.

Definition: The probability that X takes on the value x, $\Pr(X = x)$, is defined as the sum of the probabilities of all sample points in Ω that are assigned the value x. We will sometimes denote $\Pr(X = x)$ by $p(x)$ or $f(x)$. Because $p(x)$ or $f(x)$ is a function that assigns probabilities to each value x of the random variable X.

Definition: The probability distribution for a discrete variable X can be represented by a formula, a table, or a graph that provides $p(x) = \Pr(X = x)$ for all x.

EXAMPLE 2.1

Six lots of components are ready to be shipped by a supplier. The number of defective components in each lot is as follows:

Lot	1	2	3	4	5	6
Number of defectives	0	2	0	1	2	0

Please enter \offprintinfo{(Title, Edition)}{(Author)}
at the beginning of your document.

One of these lots is to be randomly selected for shipment to a customer. Let X be the number of defectives in the selected lot. The three possible X values are 0, 1, and 2. Of the six equally likely simple events, three result in $X = 0$, one in $X = 1$, and the other two in $X = 2$. Let $p(0)$ denote the probability that $X = 0$ and $p(1)$ and $p(2)$ represent the probabilities of the other two possible values of X. Then:

$$p(0) = \Pr(X = 0) = \Pr(\text{lot 1 or 3 or 6 is sent}) = \frac{3}{6} = 0.5$$
$$p(1) = \Pr(X = 1) = \Pr(\text{lot 4 is sent}) = \frac{1}{6} = 0.167$$
$$p(2) = \Pr(X = 2) = \Pr(\text{lot 2 or 5 is sent}) = \frac{2}{6} = 0.333$$

That is, a probability of 0.5 is distributed to the X value 0, a probability of 0.167 is placed on the X value 1, and the remaining probability, 0.3, is associated with the X value 2. The values of X along with their probabilities collectively specify the probability distribution or probability mass function of X. If this experiment were repeated over and over again, in the long run $X = 0$ would occur one-half of the time, $X = 1$ one-sixth of the time, and $X = 2$ one-third of the time.

EXAMPLE 2.2

A supervisor in a manufacturing plant has three men and three women working for him. He wants to choose two workers for a special job. Not wishing to show any biases in his selection, he decides to select the two workers at random. Let X denote the number of women in his selection. Find the probability distribution for X.

Solution. The supervisor can select two workers from six in $\binom{6}{2} = 15$ ways. Hence Ω contains 15 sample points, which we assume to be equally likely because random sampling was employed. Thus, $\Pr(E_i) = 1/15, i = 1, 2, \ldots, 15$. The values for X that have nonzero probability are $0, 1$, and 2. The number of ways of selecting $X = 0$ women is $\binom{3}{0}\binom{3}{2} = 1 \times 3 = 3$ sample points in the event $X = 0$, and

$$p(0) = \Pr(X = 0) = \frac{\binom{3}{0}\binom{3}{2}}{15} = \frac{3}{15} = \frac{1}{5}.$$

Similarly,

$$p(1) = \Pr(X = 1) = \frac{\binom{3}{1}\binom{3}{1}}{15} = \frac{9}{15} = \frac{3}{5}.$$
$$p(2) = \Pr(X = 2) = \frac{\binom{3}{2}\binom{3}{0}}{15} = \frac{3}{15} = \frac{1}{5}.$$

Notice that $(X = 1)$ is by far the most likely outcome. This should seem reasonable since the number of women equals the number of men in the original group. Therefore, we can write the probability function in the formula:

$$p(x) = \frac{\binom{3}{x}\binom{3}{2-x}}{\binom{6}{2}}, \quad x = 0, 1, 2.$$

Notice that, since $p(x) = \Pr(X = x)$ is a probability function, this means that the sum of $p(x)$ over the space is equal to one. One can also present the pdf of the rv X as a table as follows:

x	$p(x)$
0	1/5
1	3/5
2	1/5

◀

EXAMPLE 2.3

Consider randomly selecting a student at a large public university, and define a rv $X = 1$ if the selected student does not qualify for in-state tuition (a success from the university administration's point of view) and $X = 0$ if the student does qualify. If 20% of all students do not qualify, the pdf for X is

$$\begin{aligned} p(0) &= \Pr(X = 0) = \Pr(\text{the selected student does qualify}) = 0.8 \\ p(1) &= \Pr(X = 1) = \Pr(\text{the selected student does not qualify}) = 0.2 \\ p(x) &= \Pr(X = x) = 0 \text{ for } x \neq 0 \text{ or } 1 \end{aligned} \tag{2.1}$$

$$p(x) = \begin{cases} 0.8 & , x = 0 \\ 0.2 & , x = 1 \\ 0 & , x \neq 0, 1 \end{cases} \tag{2.2}$$

In this example, we had $p(0) = 0.8)$ and $p(1 = 0.2)$ because 20% of all students did not qualify for in-state tuition. At another university, it may be the case that $p(0) = 0.9$ and $p(1) = 0.1$. More generally, the pdf of any rv can be expressed in the form $p(1) = \alpha$ and $p(0) = 1 - \alpha$, where $0 < \alpha < 1$. Because the pdf depends on the particular value of α, we often write $p(x; \alpha)$ rather than just $p(x)$:

$$p(x; \alpha) = \begin{cases} 1 - \alpha & , x = 0 \\ \alpha & , x = 1 \\ 0 & , \text{otherwise} \end{cases} \tag{2.3}$$

EXAMPLE 2.4

Starting at a fixed time, we observe the gender of each newborn child at a certain hospital until a boy (B) is born. Let $p = \Pr(B)$, assume that successive births are independent, and define the rv X by $X =$number of births observed. Then

$$\begin{aligned} p(1) &= \Pr(X = 1) = \Pr(B) = p \\ p(2) &= \Pr(X = 2) = \Pr(GB) = \Pr(G).\Pr(B) = (1 - p)p \\ &\text{and} \\ p(3) &= \Pr(X = 3) = \Pr(GGB) = \Pr(G).\Pr(G).\Pr(B) = (1 - p)^2 p \end{aligned}$$

continuing in this way, a general formula emerges:

$$p(x) = \begin{cases} (1 - p)^{x-1} p & , x = 1, 2, 3, \ldots \\ 0 & , \text{otherwise} \end{cases} \tag{2.4}$$

The quantity p represents a number between 0 and 1 and is a parameter of the probability distribution. In the gender example, $p = 0.51$ might be appropriate.

Theorem: If $f(x)$ is a probability density function (pdf) for a discrete or continuous random variable X, then the following properties should be satisfied:

1. $f(x) \geq 0$, for all $x \in \mathcal{A}$.
2. $$1 = \begin{cases} \sum_{x\in\mathcal{A}} f(x) & \text{, discrete} \\ \\ \int_{x\in\mathcal{A}} f(x) & \text{, continuous} \end{cases}$$

For any subset of the sample points ($A \subset \mathcal{A}$), a probability set function $p(x)$ can be expressed in term of the pdf $f(x)$ as:

$$p(A) = \Pr(x \in A) = \begin{cases} \sum_{x\in A} f(x) & \text{, discrete} \\ \\ \int_{x\in A} f(x) & \text{, continuous} \end{cases}$$

EXAMPLE 2.5

Let X be a discrete r.v defined on a sample set $\mathcal{A} = \{x : x = 0, 1, 2, 3\}$, and let $f(x)$ be a function defined on $\mathcal{A}$ as: $f(x) = \frac{1}{8}\binom{3}{x}, x \in \mathcal{A}$. Examine whether $f(x)$ is a pdf of X or not. If so, find: $p(A)$ and $p(A^c)$, where $A = \{x : x = 1, 2\}$.

Solution.

- For the first condition:
$$f(x) > 0, \forall x \in \mathcal{A}.$$
- For the second condition:
$$\sum_{x\in\mathcal{A}} f(x) = \frac{1}{8}\sum_{x=0}^{3}\binom{3}{x} = \frac{1}{8}\left[\binom{3}{0} + \binom{3}{1} + \binom{3}{2} + \binom{3}{3}\right] = \frac{1}{8}(1+3+3+1) = 1.$$

This proves that $f(x)$ is a pdf of X. Hence,

$$p(A) = \sum_{x\in A} f(x) = \frac{1}{8}\sum_{x=1}^{2}\binom{3}{x} = \frac{1}{8}\left[\binom{3}{1} + \binom{3}{2}\right] = \frac{3}{4},$$

and $p(A^c) = 1 - p(A) = 1 - \frac{3}{4} = \frac{1}{4}$. ◀

EXAMPLE 2.6

Let X be a r.v defined on a sample set $\mathcal{A} = \{x : x = 1, 2, 3, \dots\}$ and let $f(x)$ be a function defined on $\mathcal{A}$ as: $f(x) = \left(\frac{1}{2}\right)^x,\ x \in \mathcal{A}$. Is $f(x)$ a pdf of X? If so, evaluate the following probabilities: $p(A), p(B), p(A \cap B), p(A \cup B)$ and $p(A|B)$, knowing that $A = \{x : x = 1, 2, 3\}$ and $B = \{x : x = 1, 3, 5, \dots\}$

Solution.

- $f(x) > 0, \quad \forall x \in \mathcal{A} = \{x : x = 1, 2, 3, \dots\}$.
- $$\sum_{x\in\mathcal{A}} f(x) = \sum_{x=1}^{\infty}\left(\frac{1}{2}\right)^x = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \cdots = \frac{1/2}{1-1/2} = 1.$$

Hence, $f(x)$ is a pdf of X.

In order to evaluate the probabilities:

- $$p(A) = \sum_{x \in A} f(x) = \sum_{x=1}^{3} \left(\frac{1}{2}\right)^x = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 = \frac{7}{8}$$

- $$p(B) = \sum_{x \in B} f(x) = \sum_{x=1,step2}^{\infty} \left(\frac{1}{2}\right)^x = \frac{1}{2} + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^5 + \cdots = \frac{1/2}{1 - 1/4} = \frac{2}{3}.$$

- $$A \cap B = \{x : x = 1, 3\} \Rightarrow p(A \cap B) = \sum_{x \in A \cap B} f(x) = \sum_{x=1,3} \left(\frac{1}{2}\right)^x = \frac{1}{2} + \left(\frac{1}{2}\right)^3 = \frac{5}{8}.$$

- $$p(A \cup B) = p(A) + p(B) - p(A \cap B) = \frac{7}{8} + \frac{2}{3} - \frac{5}{8} = \frac{11}{12}.$$

- $$p(A|B) = \frac{p(A \cap B)}{p(B)} = \frac{5/8}{2/3} = \frac{15}{16}.$$

◀

EXAMPLE 2.7

Let X be a r.v defined on a sample set $\mathcal{A} = \{x : 2 \leq x \leq 4\}$ and let $f(x)$ be a function defined on $\mathcal{A}$ as: $f(x) = \frac{1}{8}(x+1),\ x \in \mathcal{A}$. Examine whether $f(x)$ is a pdf of X? If it is, find $p(A)$, where $A = \{x : 1.5 \leq x \leq 2.5\}$.

Solution.

- $f(x) > 0, \quad \forall x \in \mathcal{A} = \{x : 2 \leq x \leq 4\}$.

- $$\int_{x \in \mathcal{A}} f(x)dx = \int_2^4 \frac{1}{8}(x+1)dx = \frac{1}{16}(x+1)^2\Big|_2^4 = \frac{1}{16}(25 - 9) = 1.$$

  Hence, $f(x)$ is a pdf of X.

- $$p(A) = \int_{x \in A} f(x)dx = \int_{1.5}^{2} f(x)dx + \int_2^{2.5} f(x)dx = 0 + \int_2^{2.5} \frac{1}{8}(x+1)dx = \frac{1}{16}(x+1)^2\Big|_2^{2.5} = 0.203.$$

◀

EXAMPLE 2.8

Suppose that the function $f(x) = e^{-x},\ x \in \mathcal{A}$ is defined as a sample set $\mathcal{A} = \{x : 0 < x < \infty\}$ and that X is a r.v. Show that $f(x)$ is a pdf of X and evaluate $p(A), p(B), p(A \cap B)$ and $p(A \cup B)$, if $A = \{x : 0 < x < 3\}$ and $B = \{x : 1 < x < \infty\}$.

Solution.

- $f(x) \geq 0, \quad \forall x \in \mathcal{A} = \{x : 0 < x < \infty\}$.
- $$\int_{x \in \mathcal{A}} f(x)dx = \int_0^\infty e^{-x}dx = -e^{-x}\Big|_0^\infty = -(e^{-\infty} - e^0) = 1.$$

  Hence, $f(x)$ is a pdf of X.

For the probability evaluations:

- $$p(A) = \int_{x \in A} f(x)dx = \int_0^3 e^{-x}dx = 1 - e^{-3}.$$
- $$p(B) = \int_{x \in B} f(x)dx = \int_1^\infty e^{-x}dx = e^{-1}.$$
- $$A \cap B = \{x : 1 < x < 3\} \Rightarrow p(A \cap B) = \int_1^3 e^{-x}dx = e^{-1} - e^{-3}.$$
- $$p(A \cup B) = p(A) + p(B) - p(A \cap B) = 1 - e^{-3} + e^{-1} - e^{-1} + e^{-3} = 1.$$

◀

EXAMPLE 2.9

Verify that the following functions are pdf's of a r.v X that defined as:

1. $f(x) = x^{-2}, \quad \mathcal{A} = \{x : 1 < x < \infty\}$.
2. $f(x) = \frac{4}{9}\binom{2}{x}\left(\frac{1}{2}\right)^x, \quad \mathcal{A} = \{x : x = 0, 1, 2\}$.
3. $f(x) = 1 - |1 - x|, \quad \mathcal{A} = \{x : 0 < x < 2\}$.
4. $f(x) = \begin{cases} 1 + x, & -1 < x < 0 \\ 1 - x, & 0 \leq x < 1 \end{cases}$

Solution.

3. $f(x) \geq 0, \forall x \in \mathcal{A} = \{x : 0 < x < 2\}$.

$$f(x) = 1 - |1 - x| = \begin{cases} 1 - (1 - x), & 1 - x \geq 0 \to x \leq 1 \\ 1 + (1 - x), & 1 - x < 0 \to x > 1 \end{cases}$$

$$f(x) = \begin{cases} x, & 0 < x \leq 1 \\ 2 - x, & 1 < x < 2 \end{cases}$$

then,

$$\int_{x\in\mathcal{A}} f(x)dx = \int_0^1 x dx + \int_1^2 (2-x)dx = \frac{1}{2}x^2\Big|_0^1 + \frac{1}{2}(2-x)^2\Big|_1^2 = \frac{1}{2} + \frac{1}{2} = 1$$

◀

EXAMPLE 2.10

Find the constant c that makes each of the following function pdf of a r.v X:

1. $f(x) = c(x+1), \quad x = 0, 1, 2, 3.$
2. $f(x) = c(x^{\alpha-1} - x^{\beta-1}), \quad 0 < x < 1, \alpha > 1, \beta > 0.$
3. $f(x) = c(1+x^2)^{-1}, \quad -\infty < x < \infty.$

Solution. Since $f(x)$ is a pdf, then it should satisfy the properties of the pdf, hence:

1.
$$1 = \sum_{x\in\mathcal{A}} f(x) = \sum_{x=0}^{3} c(x+1) = c(1+2+3+4) = 10c \Rightarrow c = 1/10.$$

2.
$$1 = \int_{x\in\mathcal{A}} f(x)dx = \int_0^1 c(x^{\alpha-1} - x^{\beta-1})dx = c\left[\frac{x^\alpha}{\alpha} - \frac{x^\beta}{\beta}\right]_0^1 = c\left(\frac{1}{\alpha} - \frac{1}{\beta}\right) = c\frac{\beta-\alpha}{\alpha\beta} \Rightarrow c = \frac{\alpha\beta}{\beta-\alpha}, \beta \neq \alpha$$

3.
$$1 = \int_{-\infty}^{\infty} c(1+x^2)^{-1}dx = c\tan^{-1}x\Big|_{-\infty}^{\infty} = c\left[\tan^{-1}(\infty) - \tan^{-1}(-\infty)\right] = c\left[\tan^{-1}(\infty) + \tan^{-1}(\infty)\right]$$

$$= 2c\tan^{-1}(\infty) = 2c\frac{\pi}{2} \Rightarrow c = \frac{1}{\pi}.$$

◀

Definition: The **Mode** of the distribution is the value of x that maximises the pdf $f(x)$ of a r.v X. Note that the mode of a continuous r.v is the solution of $f'(x) = 0$ and $f''(x) < 0$. Also, the mode may not be existed or a distribution may have more than one mode.

EXAMPLE 2.11

Find the mode for the following pdf's:

1. $f(x) = \left(\frac{1}{2}\right)^x, \quad x = 1, 2, 3, \ldots.$
2. $f(x) = 12x^2(1-x), \quad 0 < x < 1.$

Solution.

1. $f(x) = \left(\frac{1}{2}\right)^x, \quad x = 1, 2, 3, \ldots,$ then $x = 1$ is the mode.
2. $f(x) = 12x^2(1-x), \quad 0 < x < 1 \Rightarrow f'(x) = 12(2x - 3x^2)$, set $f'(x) = 0$:

$$12(2x - 3x^2) = 0 \Rightarrow x = 0, 2/3.$$

then, $f''(x) = 12(2 - 6x) = 24(1 - 3x)$.

$$f''(0) = 24(1 - 0) = 24 > 0$$

$$f''(2/3) = 24(1 - 3(2/3)) = -24 < 0$$

hence, $x = 2/3$ is the mode.

◀

2.2 Cumulative Distribution Function (CDF)

Definition: The cumulative distribution function or the probability distribution $F(x)$ of a r.v X represents the probability of $F(x) = \Pr(X \leq x)$, $-\infty < x < \infty$. Then $F(x)$ is defined on a sample space $\mathcal{A}$, as:

$$F(x) = \Pr(X \leq x) = \Pr(-\infty < X \leq x) = \begin{cases} \sum\limits_{t=-\infty}^{x} f(t), & \text{discrete.} \\ \int\limits_{t=-\infty}^{x} f(t)dt, & \text{continuous.} \end{cases}$$

EXAMPLE 2.12

A store carries flash drives with either 16, 32, 64, 128, or 256 GB of memory. The accompanying table gives the distribution of $Y =$ the amount of memory in a purchased drive:

y	16	32	64	128	256
$p(y)$	0.05	0.10	0.35	0.40	0.10

Let's first determine $F(y)$ for each of the five possible values of Y:

$$F(16) = \Pr(Y \leq 16) = \Pr(Y = 16) = p(16) = 0.05$$
$$F(32) = \Pr(Y \leq 32) = \Pr(Y = 16 \text{ or } 32) = p(16) + p(32) = 0.15$$
$$F(64) = \Pr(Y \leq 64) = \Pr(Y = 16 \text{ or } 32 \text{ or } 64) = p(16) + p(32) + p(64) = 0.50$$
$$F(128) = \Pr(Y \leq 128) = \Pr(Y = 16 \text{ or } 32 \text{ or } 64 \text{ or } 128) = p(16) + p(32) + p(64) + p(128) = 0.90$$
$$F(256) = \Pr(Y \leq 256) = \Pr(Y = 16 \text{ or } 32 \text{ or } 64 \text{ or } 128 \text{ or } 256) = p(16) + p(32) + p(64) + p(128) + p(256) = 1$$

Now for any other number y, $F(y)$ will equal the value of F at the closest possible value of Y to the left of y. For example,

$$F(43) = \Pr(Y \leq 43) = \Pr(Y \leq 32) = F(32) = 0.15$$
$$F(100) = \Pr(Y \leq 100) = \Pr(Y \leq 64) = F(64) = 0.50$$

If y is < 16, $F(y) = 0$ [e.g. $F(5) = 0$], and if y is at least 256, $F(y) = 1$ [e.g. $F(256) = 1$]. The cdf is thus

$$F(y) = \begin{cases} 0 & , y < 16 \\ 0.05 & , 16 \leq y < 32 \\ 0.15 & , 32 \leq y < 64 \\ 0.50 & , 64 \leq y < 128 \\ 0.90 & , 128 \leq y < 256 \\ 1 & , y \geq 256 \end{cases} \tag{2.5}$$

EXAMPLE 2.13

Suppose that X has a pdf $f(x)$, defined on $x = 0, 1, 2$ as:

$$f(x) = \binom{2}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{2-x}, \quad x = 0, 1, 2.$$

Find $F(x)$ for all x.

Solution. From the pdf $f(x)$, we have: $f(0) = 1/4, f(1) = 1/2, f(2) = 1/4$. In order to find the distribution function $F(x)$, from definition of $F(x) = \Pr(X \leq x)$, we should evaluate the probability for four regions of x:

$$-\infty < x < 0,\ 0 \leq x < 1,\ 1 \leq x < 2 \text{ and } 2 \leq x < \infty.$$

- For $x < 0$: Because the only value of X that are assigned positive probabilities are $0, 1$ and 2; and that non of these values are less than 0, then $f(x) = 0,\ \forall -\infty < x < 0$.
- For $0 \leq x < 1$:
$$F(0) = \Pr(0 \leq x < 1) = \Pr(X < 0) + \Pr(X = 0) = 0 + \frac{1}{4} = \frac{1}{4}.$$
- For $1 \leq x < 2$:
$$F(1) = \Pr(1 \leq x < 2) = \Pr(X = 0) + \Pr(X = 1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}.$$
- For $x \geq 2$:
$$F(2) = \Pr(X \geq 2) = \Pr(X < 0) + \Pr(0 \leq X \leq 1) + \Pr(1 \leq X < 2) = 0 + \frac{1}{4} + \frac{3}{4} = 1.$$

Therefore, the distribution probability is:

$$F(x) = \Pr(X \leq x) = \begin{cases} 0, & x < 0 \\ 1/4, & 0 \leq x < 1 \\ 3/4, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$

Notes:

1. $F(-\infty) = \lim_{x \to -\infty} F(x) = 0.$
2. $F(\infty) = \lim_{x \to \infty} F(x) = 1.$
3. $F(x)$ is a non-decreasing function of x, if $x_1 < x_2$ then $F(x_1) \leq F(x_2)$.
4. $0 \leq F(x) \leq 1$, because $0 \leq \Pr(X \leq x) \leq 1$.
5. For a continuous r.v X, the pdf could be evaluated in terms of the cdf as:
$$f(x) = \frac{d}{dx} F(x).$$
6. For a discrete r.v X, the pdf could be evaluated in terms of the cdf as:
$$f(x) = \Pr(X = x) = \Pr(X \leq x) - \Pr(X \leq x - 1) = F(x) - F(x - 1).$$

EXAMPLE 2.14

Let the r.v has a pdf $f(x) = \frac{x}{6};\ x = 1, 2, 3$. Find the cdf of X

Solution.

$$F(x) = \Pr(X \leq x) = \sum_{\tau=-\infty}^{x} f(\tau) = \sum_{\tau=1}^{x} \frac{\tau}{6} = \frac{1}{6}(1 + 2 + \cdots + x) = \frac{x(x+1)}{12}$$

$$\therefore F(x) = \begin{cases} 0, & x < 1 \\ \frac{x(x+1)}{12}, & 1 \leq x \leq 3 \\ 1, & x > 3 \end{cases}$$

EXAMPLE 2.15

Let the r.v has a pdf $f(x) = 2x^{-3};\ 1 < x < \infty$. Find cdf of X

Solution.

$$F(x) = \Pr(X \leq x) = \int_{\tau=-\infty}^{x} f(\tau) d\tau = \int_{\tau=1}^{x} 2x^{-3} d\tau = -\tau^{-2} \Big|_1^x 1 - \frac{1}{x^2}$$

$$\therefore F(x) = \begin{cases} 0, & x \leq 1 \\ 1 - \frac{1}{x^2}, & 1 < x < \infty \\ 1, & x = \infty \end{cases}$$

EXAMPLE 2.16

Let Y be a continuous r.v with probability density function given by

$$f(y) = 3y^2, \quad 0 \leq y \leq 1.$$

Find the cdf F(y).

Solution. To evaluate the cdf:

$$F(y) = \int_{-\infty}^{y} f(\tau)d\tau = \int_{0}^{y} 3\tau^2 d\tau = \tau^3|_0^y = y^3$$

therefore

$$F(y) = \begin{cases} 0 & , y < 0 \\ y^3 & , 0 \leq y \leq 1 \\ 1 & , y > 1 \end{cases}$$

◀

Remark: In order to evaluate the probability of a r.v X between a and b, we have two ways:

1. Using the pdf as:
$$\Pr(a \leq X \leq b) = \begin{cases} \sum_{x=a}^{b} f(x), & discrete \\ \int_{x=a}^{b} f(x), & continuous \end{cases}$$
2. Using the cdf as:
$$\Pr(a \leq X \leq b) = \Pr(X \leq b) - \Pr(X \leq a) = F(b) - F(a).$$

EXAMPLE 2.17

Let the r.v has a pdf $f(x) = \left(\frac{1}{2}\right)^x ;\ x = 1, 2, 3, \ldots.$

1. Find the cdf of X.
2. Evaluate $\Pr(2 \leq X \leq 4)$ using the pdf and the cdf.

Solution. 1.

$$F(x) = \sum_{\tau=-\infty}^{x} f(\tau) = \sum_{\tau=1}^{x} \left(\frac{1}{2}\right)^{\tau} = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \cdots + \left(\frac{1}{2}\right)^x = \frac{(1/2)[1-(1/2)^x]}{1-(1/2)}$$

$$\therefore F(x) = \begin{cases} 0, & x < 1 \\ 1 - \left(\frac{1}{2}\right)^x, & 1 \leq x < \infty \\ 1, & x = \infty \end{cases}$$

2. Using the pdf:
$$\Pr(2 \leq X \leq 4) = \sum_{x=2}^{4} \left(\frac{1}{2}\right)^x = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 = \frac{7}{16}.$$

Using the cdf:

$$\begin{aligned} \Pr(2 \leq X \leq 4) &= \Pr(X \leq 4) - \Pr(X < 2) = \Pr(X \leq 4) - \Pr(X \leq 1) \\ &= F(4) - F(1) = 1 - \left(\frac{1}{2}\right)^4 - 1 + \left(\frac{1}{2}\right) = \frac{1}{2} - \frac{1}{16} = \frac{7}{16} \end{aligned}$$

◀

EXAMPLE 2.18

Suppose the pdf of the magnitude X of a dynamic load on a bridge (in newtons) is given by: $f(x) = \frac{1}{8} + \frac{3}{8}x.$ for $0 \le x \le 2$. Find the cdf, and evaluate $\Pr(1 \le X \le 1.5)$.

Solution. The cdf is:

$$F(x) = \int_{-\infty}^{x} f(\tau)d\tau = \int_{0}^{x} \left[\frac{1}{8} + \frac{3}{8}\tau\right] d\tau = \frac{x}{8} + \frac{3x^2}{16}.$$

therefore

$$F(x) = \begin{cases} 0 & , x < 0 \\ \frac{x}{8} + \frac{3x^2}{16} & , 0 \le x \le 2 \\ 1 & , x > 2 \end{cases}$$

Now, to evaluate the probability,

$$\Pr(1 \le X \le 1.5) = F(1.5) - F(1) = \left[\frac{(1.5)}{8} + \frac{3(1.5)^2}{16}\right] - \left[\frac{(1)}{8} + \frac{3(1)^2}{16}\right] = \frac{19}{64} = 0.297$$

◀

EXAMPLE 2.19

The distribution of the amount of gravel (in tons) sold by a construction supply company in a given week is a continuous rv X with pdf $f(x) = \frac{3}{2}(1 - x^2)$, $0 \le x \le 1$. Find the cdf of X and $\Pr(X \le 1.5)$.

Solution. The cdf of X is obtained as:

$$F(X) = \Pr(X \le x) = \int_{-\infty}^{x} f(\tau)d\tau = \int_{0}^{x} \frac{3}{2}(1 - \tau^2)d\tau = \frac{3}{2}\left[\tau - \frac{\tau^3}{3}\right]_0^x = \frac{3}{2}\left(x - \frac{x^3}{3}\right).$$

then, the cdf can be written as:

$$F(x) = \Pr(X \le x) = \begin{cases} 0 & , x < 0 \\ \frac{3}{2}\left(x - \frac{x^3}{3}\right) & , 0 \le x \le 1 \\ 1 & , x > 1 \end{cases}$$

To calculate the probability:

$$\Pr(X \le 1.5) = F(1.5) = 1.$$

◀

Definition: The **Median** of a r.v X is the value of x for which the cdf $F(x) = \frac{1}{2}$.

EXAMPLE 2.20

Find the median of the following distributions, where their pdf's are defined as follows:

1. $f(x) = \left(\frac{1}{2}\right)^x,\ x = 1, 2, 3, \ldots.$
2. $f(x) = \frac{x}{6},\ x = 1, 2, 3.$
3. $f(x) = 3x^2,\ 0 < x < 1.$

4. $f(x) = \frac{1}{\pi}(1+x^2)^{-1},\ -\infty < x < \infty$.

Solution. 1. From ex (2.17), $F(x) = 1 - \left(\frac{1}{2}\right)^x$. To evaluate the median, set $F(x) = \frac{1}{2}$:

$$1 - \left(\frac{1}{2}\right)^x = \frac{1}{2} \Rightarrow \left(\frac{1}{2}\right)^x = \frac{1}{2} \Rightarrow x = 1.$$

The median $x = 1$.

2. From ex (2.14), $F(x) = \frac{x(x+1)}{12}$. To evaluate the median, set $F(x) = \frac{1}{2}$:

$$\frac{x(x+1)}{12} = \frac{1}{2} \Rightarrow x(x+1) = 6 \Rightarrow x^2 + x - 6 = 0 \Rightarrow x = -3, 2.$$

The median $x = 2$ since $-3 \notin \mathcal{A}$.

3. $F(x) = \int_{\tau=-\infty}^{x} f(\tau)d\tau = \int_{\tau=0}^{x} 3\tau^2 d\tau = \tau^3\Big|_0^x = x^3,\ 0 < x < 1$.
Set $F(x) = \frac{1}{2} \Rightarrow x^3 = \frac{1}{2} \Rightarrow x = \frac{1}{\sqrt[3]{2}}$ is the median.

4. $F(x) = \int_{\tau=-\infty}^{x} f(\tau)d\tau = \int_{\tau=-\infty}^{x} \frac{1}{\pi}(1+\tau)^{-1}d\tau = \frac{1}{\pi}\tan^{-1}\tau|_{-\infty}^{x} = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}x,\ -\infty < x < \infty$.
Set $F(x) = \frac{1}{2} \Rightarrow \frac{1}{2} + \frac{1}{\pi}\tan^{-1}x = \frac{1}{2} \Rightarrow x = 0$ is the median.

◀

EXAMPLE 2.21

Find the constant c in the following cdf's and find the pdf for each case.

1. X is a discrete r.v with a cdf:

$$F(x) = \begin{cases} 0, & x < 1 \\ c[1-(1/2)^x], & 1 \le x < \infty \\ 1, & x = \infty \end{cases}$$

2. X is a continuous r.v has a cdf:

$$F(x) = \begin{cases} 0, & x \le \\ cx(x+1), & 0 < x < 3 \\ 1, & x \ge 3 \end{cases}$$

Solution.

1. Since $F(x)$ is a cdf, then: $F(\infty) = 1 \Rightarrow c[1-(1/2)^\infty] = 1 \Rightarrow c(1-0) = 1 \Rightarrow c = 1$. then:

$$F(x) = \begin{cases} 0, & x < 1 \\ 1 - \left(\frac{1}{2}\right)^x, & 1 \le x < \infty \\ 1, & x = \infty \end{cases}$$

Therefore, the pdf $f(x) = F(x) - F(x-1)$:

$$f(x) = 1 - \left(\frac{1}{2}\right)^x - 1 + \left(\frac{1}{2}\right)^{x-1} = \left(\frac{1}{2}\right)^{x-1} - \left(\frac{1}{2}\right)^x = \left(\frac{1}{2}\right)^{x-1}\left(1 - \frac{1}{2}\right)$$

$$\therefore f(x) = \left(\frac{1}{2}\right)^x, \quad x = 1, 2, 3, \ldots$$

2. Since $F(\infty)$ is a cdf, then: $F(3) = 1 \Rightarrow 3c(3+1) = 1 \Rightarrow 12c = 1 \Rightarrow c = \frac{1}{2}$, then:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1}{12}x(x+1), & 0 < x < 3 \\ 1, & x \geq 3 \end{cases}$$

Therefore, the pdf $f(x) = F'(x)$:

$$\therefore f(x) = \frac{1}{12}(2x+1), \quad 0 < x < 3$$

◀

EXAMPLE 2.22

Find the cdf of the r.v X which has a pdf $f(x) = \begin{cases} x, & 0 < x < 1 \\ 2 - x, & 1 \leq x < 2 \end{cases}$

Solution.

$$F(x) = \Pr(X \leq x) = \begin{cases} 0, & x \leq 0 \\ \int_0^x f(\tau)d\tau = \int_0^x \tau d\tau = \frac{1}{2}\tau^2\Big|_0^x = \frac{1}{2}x^2, & 0 < x < 1 \\ \int_0^1 f(\tau)d\tau + \int_1^x f(\tau)d\tau = 1 - \frac{1}{2}(2-x)^2, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$

◀

2.3 Mathematical Expectation

We have observed that the probability distribution for a random variable is a theoretical model for the empirical distribution of data associated with a real population. If the model is an accurate representation of nature, the theoretical and empirical distributions are equivalent. Consequently, we attempt to find the mean and the variance for a random variable and thereby to acquire numerical descriptive measures, parameters, for the probability distribution.

Definition: Let X be a r.v with probability function $f(x)$ and $u(X)$ be a real-valued function of X. Then the expected value of $u(X)$ is given by:

$$E[u(X)] = \begin{cases} \sum_{x=-\infty}^{\infty} u(x)f(x), & \text{discrete} \\ \int_{x=-\infty}^{\infty} u(x)f(x)dx, & \text{continuous} \end{cases}$$

EXAMPLE 2.23

Let the random variable X has a pdf $f(x) = \frac{x}{6},\ x = 1, 2, 3$. Find $E[X],\ E[X^2],\ E[X^3],\ E[(X-1)^3]$.

Solution.

-
$$E[X] = \sum_x xf(x) = \sum_{x=1}^{3} x\frac{x}{6} = \frac{1}{6}(1^2 + 2^2 + 3^2) = \frac{14}{6} = \frac{7}{3}$$

-
$$E[X^2] = \sum_x x^2 f(x) = \frac{1}{6}\sum_{x=1}^{3} x^3 = \frac{1}{6}(1^3 + 2^3 + 3^3) = \frac{36}{6} = 6$$

-
$$E[X^3] = \sum_x x^3 f(x) = \frac{1}{6}\sum_{x=1}^{3} x^4 = \frac{1}{6}(1^4 + 2^4 + 3^4) = \frac{98}{6} = \frac{49}{3}$$

-
$$E[(X-1)^3] = \sum_x (x-1)^3 f(x) = \frac{1}{6}\sum_{x=1}^{3} x(x-1)^3 = \frac{1}{6}(0 + 2 + 24) = \frac{26}{6} = \frac{13}{3}$$

◄

EXAMPLE 2.24

Let the random variable X has a pdf $f(x) = \frac{1}{18}(x+2),\ -2 < x < 4$. Find $E[3X],\ E[(X+2)^3]$.

Solution.

-
$$\begin{aligned} E[3X] = \int_x 3xf(x) &= \int_{x=-2}^{4} 3x\frac{1}{18}(x+2)dx = \frac{1}{6}\int_{-2}^{4}(x^2+2x)dx = \frac{1}{6}\left[\frac{x^3}{3} + x^2\right]_{-2}^{4} \\ &= \frac{1}{6}\left[\left(\frac{64}{3} + 16\right) - \left(-\frac{8}{3} + 4\right)\right] = \frac{1}{6}\left[\frac{64}{3} + 16 = \frac{8}{3} - 4\right] = 6 \end{aligned}$$

-
$$\begin{aligned} E[(X+2)^3] &= \int_{x=-2}^{4}(x+2)^3\frac{1}{18}(x+2)dx = \frac{1}{18}\int_{-2}^{4}(x+2)^4 dx = \frac{1}{18}\frac{(x+2)^5}{5}\Big|_{-2}^{4} \\ &= \frac{1}{90}\left[6^5 - 0\right] = 86.4 \end{aligned}$$

◄

Notes:

- If c is constant, then $E[c] = c$.
- If c is constant and u is a function of X, then $E[cu(X)] = cE[u(X)]$.
- If $c_1, c_2, \ldots, c_n$ are constants, and $u_1, u_2, \ldots, u_n$ are functions, then $E\left(\sum_{i=1}^{n} c_i u_i\right) = \sum_{i=1}^{n} E[c_i u_i] = \sum_{i=1}^{n} c_i E[u_i]$.

EXAMPLE 2.25

Let the r.v X has a pdf $f(x) = 2(1-x),\ 0 < x < 1$. Find $E[X]$, $E[X^2]$ and $E[6X + 3X^2 - 4]$.

Solution. -

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = 2\int_0^1 x(1-x)dx = 2\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = \frac{1}{3}.$$

-

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = 2\int_0^1 x^2(1-x)dx = 2\left[\frac{x^3}{3} - \frac{x^4}{4}\right]_0^1 = \frac{1}{6}.$$

-

$$E[6X + 3X^2 - 4] = 6E[X] + 3E[X^2] - 4 = 6\frac{1}{3} + 3\frac{1}{6} - 4 = \frac{-3}{2}.$$

◀

2.3.1 Some Special Mathematical Expectations

In this section, we will introduce some special mathematical expectation which are most common in the use of statistical problems.

1. The **Mean** (or the expected vale) of a r.v X is the mathematical expectation $E[X]$ and denoted by μ.
2. If X is a r.v with mean $E(X) = \mu$, the **Variance** of r.v X, denoted by σ^2 or $Var(X)$, is defined to be the expected value of $(X-\mu)^2$. That is,
$$\sigma^2 = Var(x) = E[(X-\mu)^2].$$
The standard deviation of X is the positive square root of σ^2.

   Properties of Variance

   i The variance $\sigma^2 = E[(x-\mu)^2] = E[X^2 - 2\mu X + \mu^2] = E[X^2] - \mu^2$.

   ii If c is a constant, then $Var(c) = 0$.

   iii If c is a constant and X is a r.v, then $Var(cX) = c^2 Var(X)$
3. The mathematical expectation $\mu_r' = E[X^r]$ is called the rth moment about the origin.
4. The mathematical expectation $\mu_r = E[(X-\mu)^r]$ is called the rth moment about the mean.
5. The **Moment Generating Function** (mgf) of a r.v X is the expectation of $E[e^{tX}]$ (if exist), and denoted by $M(t)$. The reason of the function $M(t)$ is called mgf can be explained by the following statement. We have
$$e^{tx} = 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \dots.$$
Then, the expectation
$$\begin{aligned} E[e^{tx}] &= \sum_x e^{tx} f(x) = \sum_x \left[1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \dots\right] f(x) \\ &= \sum_x f(x) + t\sum_x xf(x) + \frac{t^2}{2!}\sum_x x^2 f(x) + \frac{t^3}{3!}\sum_x x^3 f(x) + \dots \\ &= 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \frac{t^3}{3!}\mu_3' + \dots. \end{aligned}$$

This argument involves an interchange of summations, which is justifiable if $M(t)$ exists. Thus, $E[e^{tX}]$ is a function of all the moments μ'_k about the origin, for $k = 1, 2, 3, \ldots$. In particular, μ'_k is the coefficient of $t^k/k!$ in the series expansion of $M(t)$.

Notes:

i $M(0) = 1$.

ii If we can find $E[e^{tX}]$, we can find any of the moments for X. If $M(t)$ exist, then for any positive integer k,

$$\left.\frac{d^k M(t)}{dt^k}\right|_{t=0} = M^{(k)}(0) = \mu'_k.$$

In other words, if you find the kth derivative of $M(t)$ with respect to t and then set $t = 0$, the result will be μ'_k.

iii If we set $\xi(t) = \ln M(t)$, then

$$\xi'(t) = \frac{M'(t)}{M(t)} \Rightarrow \xi'(0) = \frac{M'(0)}{M(0)} = \frac{E[X]}{1} = \mu$$

$$\xi''(t) = \frac{M(t)M''(t) - M'(t)M'(t)}{[M(t)]^2} \Rightarrow \xi''(0) = \frac{M(0)M''(0) - [M'(0)]^2}{[M(0)]^2}$$

$$\Rightarrow \xi''(0) = \frac{E[X^2] - \mu^2}{1^2} = \sigma^2.$$

EXAMPLE 2.26

The probability distribution of a r.v Y is given in the following table. Find the mean, variance and standard deviation of Y.

Probability distribution of Y

y	0	1	2	3
p(y)	1/8	1/4	3/8	1/4

Solution. By definition:

$$\mu = E[Y] = \sum_{y=0}^{3} yp(y) = 0(1/8) + 1(1/4) + 2(3/8) + 3(1/4) = 1.75$$

$$\sigma^2 = E[(Y-\mu)^2] = \sum_{y=0}^{3}(y-\mu)^2 p(y) = (0-1.75)^2(1/8)+(1-1.75)^2(1/4)+(2-1.75)^2(3/8)+(3-1.75)^2(1/4) = 0.9375$$

or

$$E[Y^2] = \sum_{y=0}^{3} y^2 p(y) = (0)^2(1/8) + (1)^2(1/4) + (2)^2(3/8) + (3)^2(1/4) = 4$$

$$\therefore \sigma^2 = E[Y^2] - \mu^2 = 4 - (1.75)^2 = 0.9375$$

and then

$$\sigma = +\sqrt{\sigma^2} = \sqrt{0.9375} = 0.97$$

◀

EXAMPLE 2.27

The manager of an industrial plant is planning to buy a new machine of either type A or type B. If t denotes the number of hours of daily operation, the number of daily repairs Y_1 required to maintain a machine of type A is a random variable with mean and variance both equal to $0.10t$. The number of daily repairs Y_2 for a machine of type B is a random variable with mean and variance both equal to $0.12t$. The daily cost of operating A is $C_A(t) = 10t + 30Y_1^2$; for B it is $C_B(t) = 8t + 30Y_2^2$. Assume that the repairs take negligible time and that each night the machines are tuned so that they operate essentially like new machines at the start of the next day. Which machine minimizes the expected daily cost if a workday consists of (a) 10 hours and (b) 20 hours?

Solution. The expected daily cost for A is

$$\begin{aligned} E[C_A(t)] &= E[10t + 30Y_1^2] = 10t + 30E[Y_1^2] \\ &= 10t + 30\{Var(Y_1) + (E[Y_1])^2\} = 10t + 30[0.10t + (0.10t)^2] \\ &= 13t + 0.3t^2. \end{aligned}$$

In this calculation, we used the known values for $Var(Y_1)$ and $E(Y_1)$ and the fact that $Var(Y_1) = E(Y_1^2) - [E(Y_1)]^2$ to obtain that $E(Y_1^2) = Var(Y_1) + [E(Y_1)]^2 = 0.10t + (0.10t)^2$. Similarly,

$$\begin{aligned} E[C_B(t)] &= E[8t + 30Y_2^2] = 8t + 30E[Y_2^2] \\ &= 8t + 30V(Y_2) + (E[Y_2])^2 = 8t + 30[0.12t + (0.12t)^2] \\ &= 11.6t + 0.432t^2. \end{aligned}$$

Thus, for scenario (a) where $t = 10$,

$$E[C_A(10)] = 160 \text{ and } E[C_B(10)] = 159.2,$$

which results in the choice of machine B.

For scenario (b), $t = 20$ and

$$E[C_A(20)] = 380 \text{ and } E[C_B(20)] = 404.8,$$

resulting in the choice of machine A.

In conclusion, machines of type B are more economical for short time periods because of their smaller hourly operating cost. For long time periods, however, machines of type A are more economical because they tend to be repaired less frequently. ◀

EXAMPLE 2.28

A retailer for a petroleum product sells a random amount X each day. Suppose that X , measured in thousands of gallons, has the probability density function $f(x) = \frac{3}{8}x^2,\ 0 \le x \le 2$. The retailer's profit turns out to be 100$ for each 1000 gallons sold if $X \le 1$ and 40$ extra per 1000 gallons if $X > 1$. Find the retailer's expected profit for any given day.

Solution. Let $p(x)$ denote the retailer's daily profit. Then

$$p(x) = \begin{cases} 100X, & 0 \le X \le 1, \\ 140X, & 1 \le X \le 2. \end{cases}$$

We want to find expected profit; by definition, the expectation is:

$$E[p(X)] = \int_x p(x)f(x)dx = \int_0^1 100x\left[\frac{3}{8}x^2\right]dx + \int_1^2 140x\left[\frac{3}{8}x^2\right]dx$$
$$= \left[\frac{300}{(8)(4)}x^4\right]_0^1 + \left[\frac{420}{(8)(4)}x^4\right]_1^2 = 206.25$$

Thus, the retailer can expect a profit of 206.25$ on the daily sale of this particular product. ◄

EXAMPLE 2.29

Find the mean and variance, if exist, for each of the following distributions, with pdf's:

1. $f(x) = \frac{x}{15}, \quad x = 1, 2, 3, 4, 5.$
2. $f(x) = \frac{1}{2}(x+1), \quad -1 < x < 1.$
3. $f(x) = x^{-2}, \quad 1 < x < \infty.$
4. $f(x) = e^{-x}, \quad 0 < x < \infty.$
5. $f(x) = \frac{3}{8}x^2, \ 0 \le x \le 2$

Solution.

1.
$$\mu = E[X] = \sum_x xf(x) = \sum_{x=1}^5 x\frac{x}{15} = \frac{1}{15}(1^2 + 2^2 + 3^2 + 4^2 + 5^2) = \frac{11}{3}.$$

$$E[X^2] = \sum_x x^2 f(x) = \sum_{x=1}^5 x^2\frac{x}{15} = \frac{1}{15}(1^3 + 2^3 + 3^3 + 4^3 + 5^3) = 15.$$

$$\sigma^2 = Var(X) = E[X^2] - \mu^2 = 15 - \left(\frac{11}{3}\right)^2 = 15 - \frac{121}{9} = \frac{14}{9}.$$

2.
$$\mu = E[X] = \int_x xf(x)dx = \int_{-1}^1 x\frac{1}{2}(x+1)dx = \frac{1}{2}\left[\frac{x^3}{3} + \frac{x^2}{2}\right]_{-1}^1 = \frac{1}{2}\left[\left(\frac{1}{3} + \frac{1}{2}\right) - \left(\frac{-1}{3} + \frac{1}{2}\right)\right] = \frac{1}{3}.$$

$$E[X^2] = \int_x x^2 f(x)dx = \int_{-1}^1 x\frac{1}{2}(x+1)dx = \frac{1}{2}\left[\frac{x^4}{4} + \frac{x^3}{3}\right]_{-1}^1 = \frac{1}{3}.$$

$$\sigma^2 = Var(X) = E[X^2] - \mu^2 = \frac{1}{3} - \frac{1}{9} = \frac{2}{9}.$$

3.
$$\mu = E[X] = \int_x xf(x)dx = \int_1^\infty xx^{-2}dx = \int_1^\infty \frac{1}{x}dx = \ln x\Big|_1^\infty = \ln\infty - \ln 1 = \infty - 0 = \infty.$$

Therefore, the mean μ does not exist, hence the variance σ^2 does not exist neither.

4.
$$\mu = E[X] = \int_x xf(x)dx = \int_0^\infty xe^{-x}dx = -xe^{-x} - e^{-x}\Big|_0^\infty = (0-0) - (0-1) = 1.$$
$$E[X^2] = \int_x x^2 f(x)dx = \int_0^\infty xe^{-x}dx = -x^2e^{-x} - 2xe^{-x} - 2e^{-x}\Big|_0^\infty = 2.$$
$$\sigma^2 = E[X^2] - \mu^2 = 2 - (1)^2 = 1.$$

5.
$$\mu = E[X] = \int_0^2 xf(x)dx = \int_0^2 x\frac{3}{8}x^2dx = \frac{3}{8}\frac{x^4}{4}\Big|_0^2 = 1.5.$$
$$E[X^2] = \int_0^2 x^2 f(x)dx = \int_0^2 x^2\frac{3}{8}x^2dx = \frac{3}{8}\frac{x^5}{5}\Big|_0^2 = 2.4$$
$$\sigma^2 = Var(X) = E[X^2] - (E[X])^2 = 2.4 - (1.5)^2 = 0.15$$

◀

EXAMPLE 2.30

Let the r.v X has a pdf $f(x) = \left(\frac{1}{2}\right)^x,\ x = 1, 2, 3, \ldots$, then:

1. Find the mgf of X.
2. Evaluate the mean and variance of X using the mgf.

Solution.

1. The mgf $M(t)$ is the expectation of the function e^{tX}, then
$$M(t) = E[e^{tX}] = \sum_x e^{tx}f(x) = \sum_{x=1}^\infty e^{tx}\left(\frac{1}{2}\right)^x = \sum_{x=1}^\infty \left(\frac{e^t}{2}\right)^x = \frac{e^t}{2} + \left(\frac{e^t}{2}\right)^2 + \left(\frac{e^t}{2}\right)^3 + \cdots = \frac{(e^t/2)}{1-(e^t/2)}$$
$$\therefore M(t) = \frac{e^t}{2-e^t}, \quad t \neq \ln 2.$$

2. Check $M(0) = \frac{e^0}{2-e^0} = \frac{1}{2-1} = 1$. now
$$M'(t) = \frac{(2-e^t)e^t - e^t(-e^t)}{(2-e^t)^2} = \frac{2e^t - e^{2t} + e^{2t}}{(2-e^t)^2} = \frac{2e^t}{(2-e^t)^2}$$
$$M'(0) = \frac{2e^0}{(2-e^0)^2} = \frac{2}{(2-1)^2} = 2 = \mu$$
and that
$$M''(t) = \frac{(2-e^t)^2 2e^t - 2e^t 2(2-e^t)(-e^t)}{(2-e^t)^4} = \frac{(2-e^t)2e^t + 4e^{2t}}{(2-e^t)^3}$$
$$M''(0) = \frac{(2-1)2+4}{(2-1)^3} = 6 = E[X^2]$$

then

$$\sigma^2 = E[X^2] - \mu^2 = 6 - 4 = 2.$$

or, we consider the function $\xi(t) = \ln M(t) = t - \ln(2 - e^t)$, then

$$\xi'(t) = 1 + \frac{e^t}{2 - e^t} \Rightarrow \xi'(0) = 1 + \frac{1}{2-1} = 2 = \mu$$

and

$$\xi''(t) = \frac{(2 - e^t)e^t - e^t(-e^t)}{(2 - e^t)^2} = \frac{2e^t}{(2 - e^t)^2} \Rightarrow \xi''(0) = \frac{2}{(2-1)^2} = 2 = \sigma^2.$$

◀

EXAMPLE 2.31

Find the mgf of a r.v X that has a pdf $f(x) = xe^{-x},\ 0 < x < \infty$, then evaluate the mean and variance of X.

Solution. The mgf od r.v X is the expectation of e^{tX},

$$M(t) = E[e^{tX}] = \int_x e^{tx} f(x)dx = \int_0^\infty e^{tx} xe^{-x} dx = \int_0^\infty xe^{-(1-t)x} dx = \left[-\frac{xe^{-(1-t)x}}{1-t} - \frac{e^{-(1-t)x}}{(1-t)^2} \right]_0^\infty$$

Therefore,

$$M(t) = \frac{1}{(1-t)^2},\ t \neq 1.$$

In order to evaluate μ and σ^2, we have the mgf $M(t) = (1-t)^{-2}$,

$$M'(t) = 2(1-t)^{-3} \Rightarrow M'(0) = 2 = \mu$$

$$M''(t) = 6(1-t)^{-4} \Rightarrow M''(0) = 6 = E[X^2]$$

$$\therefore \sigma^2 = E[X^2] - \mu^2 = 6 - 4 = 2$$

or, we consider the function $\xi(t) = \ln M(t) = -2\ln(1-t)$, then

$$\xi'(t) = \frac{2}{1-t} \Rightarrow \xi'(0) = 2 = \mu$$

$$\xi''(t) = 2(1-t)^{-2} \Rightarrow \xi''(0) = 2 = \sigma^2$$

◀

EXAMPLE 2.32

A manufacturing company ships its product in two different sizes of truck trailers. Each shipment is made in a trailer with dimensions 8 feet $\times$ 10 feet $\times$ 30 feet or 8 feet $\times$ 10 feet $\times$ 40 feet. If 30% of its shipments are made by using 30-foot trailers and 70% by using 40-foot trailers, find the mean volume shipped per trailer load. (Assume that the trailers are always full.)

Solution. Assume that the volume of the 30-foot trailers is v_1 and the 40-foot trailers is v_2, then:

$$v_1 = 8 \times 10 \times 30 = 2400 \text{ feet}^3.$$

$$v_2 = 8 \times 10 \times 40 = 3200 \text{ feet}^3.$$

since we have the probability of shipping throughout v_1 and v_2 are:

$$p(v_1) = 30\% = \frac{3}{10} \quad p(v_2) = 70\% = \frac{7}{10}.$$

Therefore, the expected shipping volume is:

$$E[V] = \sum_{i=1}^{2} v_i p(v_i) = 2400 \times \frac{3}{10} + 3200 \times \frac{7}{10} = 2990 \text{ feet}^3.$$

◀

EXAMPLE 2.33

In a gambling game a person draws a single card from an ordinary 52-card playing deck. A person is paid 15$ for drawing a jack or a queen and 5$ for drawing a king or an ace. A person who draws any other card pays 4$. If a person plays this game, what is the expected gain?

Solution. Let the r.v X represents the outcome of the draw. Then, the player gain could be represented as:

$$g = \begin{cases} 15, & x = J, Q \\ 5, & x = K, A \\ -4, & x = 2, 3, 4, 5, 6, 7, 8, 9, 10 \end{cases}$$

Since we have $\binom{52}{4} = \frac{1}{13}$ ways of drawing a card, then the probability of drawing any number or shape is equal to $\frac{1}{13}$, i.e:

Probability distribution of X

x	2	3	...	10	J	Q	K	A
p(x)	1/13	1/13	...	1/13	1/13	1/13	1/13	1/13

Then, the expected gain of the played is calculated by:

$$E[G] = \sum gp(x) = \left[9\left((-4) \times \frac{1}{13}\right) + 2\left((5) \times \frac{1}{13}\right) + 2\left((15) \times \frac{1}{13}\right)\right] = \frac{-36}{13} + \frac{10}{13} + \frac{30}{13} = \frac{4}{13} = 0.307.$$

◀

EXAMPLE 2.34

A builder of houses needs to order some supplies that have a waiting time Y for delivery, with a continuous uniform distribution over the interval from 1 to 4 days $\left(p(y) = \frac{1}{3},\ 1 \leq y \leq 4\right)$. Because he can get by without them for 2 days, the cost of the delay is fixed at 100$ for any waiting time up to 2 days. After 2 days, however, the cost of the delay is 100$ + 20$ per day (prorated) for each additional day. That is, if the waiting time is 3.5 days, the cost of the delay is $100\$ + 20\$(1.5) = 130\$$. Find the expected value of the builder's cost due to waiting for supplies.

Solution. Assume that the cost of waiting the supplies W_c, and Y is the r.v that represents the number of waiting days, then:

$$W_c = \begin{cases} 100, & 1 \leq y \leq 2 \\ 100 + 20(y-2), & 2 < y \leq 4 \end{cases}$$

Therefore, expected value of the builder's cost due to waiting for supplies is

$$E[W_c] = \int W_c p(y)dy = \int_1^2 100\frac{1}{3}dy + \int_2^4 (100 + 20(y-2))\frac{1}{3}dy$$
$$= \frac{100}{3}y\Big|_1^2 + \frac{100}{3}y\Big|_2^4 + \frac{20}{3}\left[\frac{y^2}{2} - 2y\right]_2^4 = \frac{100}{3} + \frac{200}{3} + \frac{40}{3} = \frac{340}{3} = 113.33$$

◀

2.3.2 Tchebyshev's Inequality

In order to find the upper and lower bounds for certain probability, we will need to prove some theorems. These bounds are not necessarily close to the exact probability.

***Theorem*:** Let $u(X)$ be a non-negative function of a r.v X whose pdf $f(x),\ -\infty < x < \infty$. If $E[u(X)]$ exist, then for all positive constant c,

$$\Pr[u(X) \geq c] \leq \frac{E[u(X)]}{c}.$$

***Theorem*:** *Tchebyshev Inequality*: Let X be a r.v with mean μ and finite variance σ^2. Then, for any constant $k > 0$,

$$\Pr\left(|X - \mu| < k\sigma\right) \geq 1 - \frac{1}{k^2}, \text{ or } \Pr\left(|X - \mu| \geq k\sigma\right) \leq \frac{1}{k^2}$$

Two important aspects of this result should be pointed out. First, the result applies for any probability distribution. Second, the results of the theorem are very conservative in the sense that the actual probability that X is in the interval $\mu \pm k\sigma$ usually exceeds the lower bound for the probability, $1 - 1/k^2$, by a considerable amount.

***Proof*:** Consider the previous theorem by taking $u(X) = (X - \mu)^2$ and $c^2 = k^2\sigma^2$, then

$$\Pr\left[(X - \mu)^2 \geq k^2\sigma^2\right] \leq \frac{E\left[(X - \mu)^2\right]}{k^2\sigma^2} = \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

Since $(x - \mu)^2 \geq k^2\sigma^2 \Leftrightarrow |x - \mu| \geq k\sigma$. It follows,

$$\Pr\left(|X - \mu| \geq k\sigma\right) \leq \frac{1}{k^2}$$

EXAMPLE 2.35

Let the r.v X has a pdf $f(x) = \frac{1}{2\sqrt{3}},\ -\sqrt{3} < x < \sqrt{3}$. Find The exact value of $\Pr\left(|X - \mu| \geq \frac{3}{2}\sigma\right)$ and $\Pr\left(|X - \mu| \geq 2\sigma\right)$, and then compare those results with their upper bounds.

Solution. First of all, we need to find the mean μ and the variance σ^2. Then

$$\mu = E[X] = \int_x xf(x)dx = \frac{1}{2\sqrt{3}}\int_{-\sqrt{3}}^{\sqrt{3}} xdx = 0.$$

$$E[X^2] = \int_x x^2 f(x)dx = \frac{1}{2\sqrt{3}}\int_{-\sqrt{3}}^{\sqrt{3}} x^2dx = 1.$$

$$\sigma^2 = E[X^2] - \mu^2 = 1 \Rightarrow \sigma = 1.$$

The exact value of probability

$$\Pr\big(|X-\mu|\geq\frac{3}{2}\sigma\big)=\Pr\big(|X|\geq\frac{3}{2}\big)=1-\Pr\big(|X|<\frac{3}{2}\big)=1-\Pr\big(-\frac{3}{2}<X<\frac{3}{2}\big)$$
$$=1-\int_{-3/2}^{3/2}f(x)dx=1-\frac{1}{2\sqrt{3}}\int_{-3/2}^{3/2}dx=1-\frac{\sqrt{3}}{2}=0.134$$

To compare with the probability upper bound, we will use Tchebyshev inequality, to find this upper bound for probability $\Pr\Big(|X-\mu|\geq\frac{3}{2}\sigma\Big)$

$$\Pr\Big(|X|\geq\frac{3}{2}\Big)\leq\left(\frac{2}{3}\right)^2=\frac{4}{9}=0.44.$$

It is clear that the exact probability (0.134) is less than the upper bound (0.44).

For the next part, we do the same. The exact value of probability

$$\Pr\big(|X-\mu|\geq 2\sigma\big)=\Pr\big(|X|\geq 2\big)=1-\Pr\big(|X|<2\big)=1-\Pr\big(-2<X<2\big)$$
$$=1-\int_{-2}^{2}f(x)dx=1-\left[\int_{-2}^{\sqrt{3}}f(x)dx+\int_{-\sqrt{3}}^{\sqrt{3}}f(x)dx+\int_{\sqrt{3}}^{2}f(x)dx\right]=1-(0+1+0)=0$$

To compare with the probability upper bound, we will use Tchebyshev inequality, to find this upper bound for probability $\Pr\big(|X-\mu|\geq 2\sigma\big)$

$$\Pr\big(|X|\geq 2\big)\leq\frac{1}{2^2}=0.25.$$

It is clear that the exact probability (0) is less than the upper bound (0.25). ◀

***Note*:** We may have the mean μ and variance σ^2 for a distribution whose pdf is not available for some reason. In this case, to find a certain probability, we use Tchebyshev inequality to find the upper or lower bound for this probability.

EXAMPLE 2.36

Let the r.v X has mean $\mu=3$ and variance $\sigma^2=4$. Use Tchebyshev inequality to determine a lower bound for $\Pr(-2<X<8)$.

Solution. To use the Tchebyshev inequality, we need to get to the form $\Pr\big[|X-\mu|<k\sigma\big]=1-\frac{1}{k^2}$. Then

$$\Pr\big(-2<X<8\big)=\Pr\big(-2-3<X-\mu<8-3\big)=\Pr\big(-5<X-\mu<5\big)=\Pr\big(|X-\mu|<5\big)$$
$$=\Pr\big(|X-\mu|<\frac{5}{2}\sigma\big)\geq 1-\left(\frac{2}{5}\right)^2=1-\frac{4}{25}=\frac{21}{25}=0.85$$

◀

EXAMPLE 2.37

The number of customers per day at a sales counter, Y , has been observed for a long period of time and found to have mean 20 and standard deviation 2. The probability distribution of Y is not known. What can be said about the probability that, tomorrow, Y will be greater than 16 but less than 24?

Solution. We want to find $\Pr(16<Y<24)$. From Tchebyshev inequality we know, for any $k\geq 0$, $\Pr(|Y-\mu|<k\sigma)\geq 1-1/k^2$, or

$$\Pr\big[(\mu-k\sigma)<Y<(\mu+k\sigma)\big]\geq 1-\frac{1}{k^2}.$$

Because $\mu = 20$ and $\sigma = 2$, it follows that $\mu - k\sigma = 16$ and $\mu + k\sigma = 24$ if $k = 2$. Thus

$$\Pr\big(16 < Y < 24\big) = \Pr\big(\mu - 2\sigma < Y < \mu + 2\sigma\big) \geq 1 - \frac{1}{2^2} = \frac{3}{4}.$$

In other words, tomorrow's customer total will be between 16 and 24 with a fairly high probability (at least $3/4$).

Notice that if σ were 1, k would be 4, and

$$\Pr\big(16 < Y < 24\big) = \Pr\big(\mu - 4\sigma < Y < \mu + 4\sigma\big) \geq 1 - \frac{1}{4^2} = \frac{15}{16}.$$

Thus, the value of σ has considerable effect on probabilities associated with intervals. ◀

EXAMPLE 2.38

Suppose that experience has shown that the length of time T (in minutes) required to conduct a periodic maintenance check on a dictating machine follows a gamma distribution with mean $\mu = 6.2$ and variance $\sigma^2 = 12.4$. A new maintenance worker takes 22.5 minutes to check the machine. Does this length of time to perform a maintenance check disagree with prior experience?

Solution. We know that $\mu = 6.2$ and $\sigma^2 = 12.4 \Rightarrow \sigma = \sqrt{12.4} = 3.52$. We need to evaluate $\Pr(T \geq 22.5)$, then

$$\Pr\big(T - \mu \geq 22.5 - \mu\big)$$

. Notice that $t = 22.5$ minutes exceeds the mean $\mu = 6.2$ minutes by 16.3 minutes, or $k = 16.3/3.52 = 4.63$ standard deviations. Then from Tchebysheff's theorem,

$$\Pr\big(|T - 6.2| \geq 16.3\big) = \Pr\big(|T - \mu| \geq 4.63\sigma\big) \leq \frac{1}{(4.63)^2} = 0.0466.$$

This probability is based on the assumption that the distribution of maintenance times has not changed from prior experience. Then, observing that $\Pr(T \geq 22.5)$ is small, we must conclude either that our new maintenance worker has generated by chance a lengthy maintenance time that occurs with low probability or that the new worker is somewhat slower than preceding ones. Considering the low probability for $\Pr(T \geq 22.5)$, we favour the latter view. ◀

2.4 Transformation of Variables (cdf technique)

Assume that X is a r.v defined on sample space $\mathcal{A}$ and has pdf $f(x)$ and cdf $F(x)$. Consider a new r.v Y as a function of X, say $Y = \psi(X)$ defined on a sample space $\mathcal{B} = \{y : y = \psi(x),\ x \in \mathcal{A}\}$. The aim now is to find the distribution of r.v Y, let $G(y) = \Pr(Y \leq y)$ and $g(x)$ represent the cdf and the pdf of Y. If the function $y = \psi(x)$ is a one-to-one transformation that maps the space $\mathcal{A}$ on to the space $\mathcal{B}$, $\Big(y = \psi(x) : \mathcal{A} \xrightarrow[on\ to]{1-1} \mathcal{B}\Big)$, then the inverse function $x = \psi^{-1}(y)$ exist. Therefore the cdf of Y can be written as:

$$G(y) = \Pr(Y \leq y) = \Pr(\psi(X) \leq y) = \Pr(X \leq \psi^{-1}(y)) = F\big(\psi^{-1}(y)\big), \quad y \in \mathcal{B}.$$

Hence, the pdf $g(y)$ is:

$$g(y) = \begin{cases} G(y) - G(y-1), & \text{discrete} \\ G'(y), & \text{continuous} \end{cases}$$

EXAMPLE 2.39

Let the r.v X has a pdf $f(x) = \frac{1}{2},\ -1 < x < 1$. Find the distribution of r.v $Y = X^2$.

Solution. Case 1 Since we have the pdf $f(x) = \frac{1}{2},\ -1 < x < 1$, then:

$$F(x) = \Pr(X \leq x) = \int_{-1}^{x} f(\tau)d\tau = \begin{cases} 0, & x \leq -1 \\ \frac{1}{2}(x+1), & -1 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

Let $G(y)$ and $g(y)$ represent the cdf and the pdf of Y defined on sample space $\mathcal{B} = \{y : 0 < y < 1\}$, and

$$\begin{aligned} G(y) &= \Pr(Y \leq y) = \Pr(X^2 \leq y) = \Pr(-\sqrt{y} \leq X \leq \sqrt{y}) = \Pr(X \leq \sqrt{y}) - \Pr(X \leq -\sqrt{y}) \\ &= F(\sqrt{y}) - F(-\sqrt{y}) = \frac{1}{2}(\sqrt{y}+1) - \frac{1}{2}(-\sqrt{y}+1) = \frac{1}{2}2\sqrt{y} \end{aligned}$$

$$\therefore G(y) = \begin{cases} 0, & y \leq 0 \\ \sqrt{y}, & 0 < y < 1 \\ 1, & y \geq 1 \end{cases}$$

Then, the pdf $g(y) = G'(y) = \frac{1}{2\sqrt{y}},\ 0 < y < 1$.

Case 2 The pdf of r.v X $f(x) = \frac{1}{2},\ -1 < x < 1$. Let $G(y)$ and $g(y)$ be the cdf and pdf of $Y = X^2$ with sample space $\mathcal{B} = \{y : 0 < y < 1\}$. Then:

$$G(y) = \Pr(Y \leq y) = \Pr(X^2 \leq y) = \Pr(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)dx = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2}dx = \sqrt{y}$$

$$\therefore G(y) = \begin{cases} 0, & y \leq 0 \\ \sqrt{y}, & 0 < y < 1 \\ 1, & y \geq 1 \end{cases}$$

◀

EXAMPLE 2.40

The function $f(x) = \frac{x}{6},\ x = 1, 2, 3$ is a pdf of r.v X that is defined on sample space $\mathcal{A} = \{1, 2, 3\}$. Find the cdf of r.v $Y = 2X + 1$.

Solution. The cdf of r.v X is:

$$F(x) = \Pr(X \leq x) = \sum_{\tau=1}^{x} f(\tau) = \frac{1}{6}\sum_{\tau=1}^{x} \tau = \frac{1}{6}(1 + 2 + \cdots + x)$$

$$\therefore F(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{12}x(x+1), & 1 \leq x \leq 3 \\ 1, & x > 3 \end{cases}$$

Let $G(y)$ is the cdf of r.v $Y = 2X + 1$ defined on sample space $\mathcal{B} = \{3.5.7\}$, then:

$$G(y) = \Pr(Y \leq y) = \Pr(2X + 1 \leq y) = \Pr\left(X \leq \frac{1}{2}(y-1)\right) = F\left(\frac{y-1}{2}\right)$$
$$= \frac{1}{12}\left(\frac{y-1}{2}\right)\left(\frac{y-1}{2}+1\right) = \frac{1}{12}\left(\frac{y-1}{2}\right)\left(\frac{y-1+2}{2}\right) = \frac{1}{48}(y-1)(y+1)$$

$$\therefore G(y) = \begin{cases} 0, & y < 3 \\ \frac{1}{48}\left(y^2-1\right), & 3 \leq y \leq 7 \\ 1, & y > 7 \end{cases}$$

◀

EXAMPLE 2.41

Find the distribution of the r.v $Y = -\ln X$, where X is defined on sample space $\mathcal{A} = \{x : 0 < x < 1\}$ with pdf $f(x) = 1,\ x \in \mathcal{A}$.

Solution. In order to evaluate the cdf of r.v X:

$$F(x) = \Pr(X \leq x) = \int_0^t f(\tau)d\tau = \int_0^x d\tau = \tau\Big|_0^x$$

$$\therefore F(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

Assume $G(y)$ and $g(y)$ represent the cdf and pdf of r.v $Y = -\ln X$ defined on sample space $\mathcal{B} = \{y : 0 < y < \infty\}$,

$$G(y) = \Pr(Y \leq y) = \Pr(-\ln X \leq y) = \Pr(\ln X > -y) = \Pr(X \geq e^{-y})$$
$$= 1 - \Pr(X \leq e^{-y}) = 1 - F\left(e^{-y}\right)$$

$$\therefore G(y) = \begin{cases} 0, & y \leq 0 \\ 1 - e^{-y}, & 0 < y < \infty \\ 1, & y = \infty \end{cases}$$

This leads to the pdf of r.v Y, $g(y) = G'(y) = e^{-y}, y \in \mathcal{B}$ ◀

EXAMPLE 2.42

A process for refining sugar yields up to 1 ton of pure sugar per day, but the actual amount produced, Y , is a random variable because of machine breakdowns and other slowdowns. Suppose that Y has density function $f(y) = 2y, 0 < y < 1$. The company is paid at the rate of \$300 per ton for the refined sugar, but it also has a fixed overhead cost of \$100 per day. Thus the daily profit, in hundreds of dollars, is $U = 3Y - 1$. Find the probability density function for U.

Solution. The function of interest here is $U = h(y) = 3y - 1$, which is increasing in y. In order to evaluate the cdf of r.v Y:

$$F(y) = \Pr(Y \leq y) = \int_{t=-\infty}^{y} f(t)dt = \int_{t=0}^{y} 2tdt = y^2.$$

Therefore

$$F(y) = \begin{cases} 0, & y \leq 0 \\ y^2, & 0 < y < 1 \\ 1, & y \geq 1 \end{cases}$$

Assume $G(u)$ and $g(u)$ represent the cdf and pdf of r.v $U = 3y - 1$ defined on sample space $\mathcal{B} = \{u : -1 < u < 2\}$,

$$G(u) = \Pr(U \leq u) = \Pr(3Y - 1 \leq u) = \Pr\left(Y \leq \frac{y+1}{3}\right) = F\left(\frac{y+1}{3}\right)$$

$$\therefore G(u) = \begin{cases} 0, & y \leq -1 \\ \frac{(y+1)^2}{9}, & -1 < u < 2 \\ 1, & u \geq 2 \end{cases}$$

This leads to the pdf of r.v U, $g(u) = G'(u) = \frac{2(y+1)}{9}, u \in \mathcal{B}$ ◀

CHAPTER 3

SOME SPECIAL MATHEMATICAL DISTRIBUTIONS

As stated in Chapter 1, a random variable is a real-valued function defined over a sample space. Consequently, a random variable can be used to identify numerical events that are of interest in an experiment. For any r.v, we can define many distribution functions in order to be able to calculate probabilities for certain events. In this chapter, we will introduce some special probability distribution for some discrete and continuous r.v's.

3.1 Discrete Distributions

In this section, some of the most important and popular distributions for discrete r.v are presented. Both the pdf and cdf are derived, and some important properties and mathematical expectation for these distribution are obtained.

3.1.1 Binomial Distribution

Some experiments consist of the observation of a sequence of identical and independent trials, each of which can result in one of two outcomes. For instance, each item leaving a manufacturing production line is either defective or non-defective. Each shot in a sequence of firings at a target can result in a hit or a miss, and each of n persons questioned prior to a local election either favors candidate Jones or does not. In this section we are concerned with experiments, known as binomial experiments, that exhibit the following characteristics.

***Definition*:** A binomial experiment possesses the following properties:

1. The experiment consists of a fixed number, n, of identical trials.
2. Each trial results in one of two outcomes: success, S, or failure, F.

Please enter \offprintinfo{(Title, Edition)}{(Author)}
at the beginning of your document.

3. The probability of success on a single trial is equal to some value p and remains the same from trial to trial. The probability of a failure is equal to $q = (1 - p)$.

4. The trials are independent.

5. The random variable of interest, X, the number of successes observed during the n trials.

Determining whether a particular experiment is a binomial experiment requires examining the experiment for each of the characteristics just listed. Notice that the random variable of interest is the number of successes observed in the n trials. It is important to realize that a success is not necessarily "good" in the everyday sense of the word. In our discussions, success is merely a name for one of the two possible outcomes on a single trial of an experiment.

***Definition*:** A random variable X is said to have a binomial distribution based on n trials with success probability p if and only if

$$f(x) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \ldots, n \text{ and } 0 \leq p \leq 1,$$

and is denoted by $X \sim B(n, p)$.

The term binomial experiment derives from the fact each trial results in one of two possible outcomes and that the probabilities $p(y), y = 0, 1, 2, \ldots, n$, are terms of the binomial expansion

$$(p+q)^n = \binom{n}{0} p^n q^0 + \binom{n}{1} p^{n-1} q^1 + \binom{n}{2} p^{n-2} q^2 + \cdots + \binom{n}{n} p^0 q^n.$$

Now, in order to verify that $f(x)$ is a valid pdf, one can easily prove that

1. $f(x) > 0, \ \forall x \in \mathcal{A} = \{x : x = 0, 1, 2, \ldots, n\}$.

2. Since $f(x)$ satisfies the binomial expansion, and that $p + q = 1$, then

$$\sum_x f(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (p+q)^n = 1^n = 1.$$

The binomial probability distribution has many applications because the binomial experiment occurs in sampling for defectives in industrial quality control, in the sampling of consumer preference or voting populations, and in many other physical situations. We will illustrate with a few examples.

EXAMPLE 3.1

Suppose that a lot of 5000 electrical fuses contains 5% defectives. If a sample of 5 fuses is tested, find the probability of observing at least one defective.

Solution. It is reasonable to assume that X, the number of defectives observed, has an approximate binomial distribution with $p = 0.05$, and $q = 0.95$. Thus,

$$\begin{aligned} \Pr(\text{at least one defective}) &= 1 - f(0) = 1 - \binom{5}{0} p^0 q^5 \\ &= 1 - (0.95)^5 = 1 - 0.774 = 0.226 \end{aligned}$$

Notice that there is a fairly large chance of seeing at least one defective, even though the sample is quite small. ◀

***The Cumulative Distribution Function*:** The cdf of X that has a binomial pdf is defined as:

$$F(x) = \Pr(X \leq x) = \sum_{\tau=0}^{x} f(\tau) = \sum_{\tau=0}^{x} \binom{n}{\tau} p^\tau q^{n-\tau}.$$

In practice, it is not easy or convenient to use the above form of the cfd to calculate the probability at certain point $F(x)$. Instead, we use Table 1 (p: 839-841)

EXAMPLE 3.2

The large lot of electrical fuses of the last example is supposed to contain only 5% defectives. If $n = 20$ fuses are randomly sampled from this lot, find the probability that at least four defectives will be observed.

Solution. Letting X denote the number of defectives in the sample, we assume the binomial model for X , with $p = 0.05$. Thus,

$$\Pr(X \geq 4) = 1 - \Pr(X \leq 3),$$

and using Table 1, we obtain

$$\Pr(X \leq 3) = \sum_{x=0}^{3} f(x) = 0.984$$

The value 0.984 is found in the table labelled $n = 20$ in Table 1. Then, the probability of getting at least 4 defective fuses is

$$\Pr(X \geq 4) = 1 - 0.984 = 0.016.$$

This probability is quite small. If we did indeed observe more than three defectives out of 20 fuses, we might suspect that the reported 5% defect rate is erroneous. ◄

***The Moment Generating Function*:** The mgf of X could be evaluated as:

$$\begin{aligned} M(t) = E\left[e^{tX}\right] &= \sum_{x} e^{tx} f(x) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} \left(pe^t\right)^x q^{n-x} \\ &= \left(pe^t + q\right)^n, \quad -\infty < t < \infty \end{aligned}$$

***Mean and Variance*:** Let X be a binomial random variable based on n trials and success probability p. Then

$$\mu = E[X] = np \quad \text{and } \sigma^2 = Var(X) = npq.$$

The derivation of the mean and variance could be done in three different ways; (1) by the direct approach of $E[X]$, (2) differentiating the mgf $M(t)$, or (3) the differentiation of the function $\ln(M(t))$. We will use the second way to evaluate the mean and variance as:

$$M'(t) = npe^t \left(pe^t + q\right)^{n-1} \Rightarrow M'(0) = np = E[X] = \mu.$$

$$M''(t) = np\left[(n-1)p\, e^{2t} \left(pe^t + q\right)^{n-2} + e^t \left(pe^t + q\right)^{n-1}\right]$$

$$M''(0) = np[(n-1)p + 1] = np(np - p + 1) = np(np + q) = n^2p^2 + npq = E[X^2]$$

$$\therefore Var(X) = \sigma^2 = E[X^2] - \mu^2 = n^2p^2 + npq - n^2p^2 = npq$$

EXAMPLE 3.3

Let the r.v $X \sim b(7, 0.5)$. Write down the pdf and mgf of X. Then find μ and σ^2. Evaluate $\Pr(X \leq 2)$, $\Pr(3 \leq X \leq 5)$, $\Pr(X = 5)$.

Solution.

- Since $X \sim b(7, 0.5)$, then the pdf of X

$$f(x) = \Pr(X = x) = \binom{7}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{7-x} = \binom{7}{x}\left(\frac{1}{2}\right)^7, \quad x = 0, 1, 2, \ldots, 7$$

- The mgf of X

$$M(t) = \left(\frac{1}{2}e^t + \frac{1}{2}\right)^7$$

- Mean and Variance: $\mu = np = \frac{7}{2}$, and $\sigma^2 = npq = \frac{7}{4}$.
- $\Pr(X \leq 2) = 0.2266$
- $\Pr(3 \leq X \leq 5) = \Pr(X \leq 5) - \Pr(X \leq 3) = \Pr(X \leq 5) - \Pr(X \leq 2) = 0.9375 - 0.2266 = 0.7109.$
- $\Pr(X = 5) = \Pr(X \leq 5) - \Pr(X \leq 4) = 0.9375 - 0.7734 = 0.1641.$

◀

EXAMPLE 3.4

A die is tossed 5 times. What is the probability of obtaining exactly three two's?

Solution. The probability of getting two's from a die tossing is $p = \frac{1}{6}$. Then, the probability of getting anything else is $q = \frac{5}{6}$. If we assume that a r.v X represents the number of two's in 5 tosses, then $X \sim b(5, \frac{1}{6})$ with pdf $f(x) = \binom{5}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{5-x}$, $x = 0, 1, 2, 3, 4, 5$. Therefore:

$$\Pr(X = 3) = f(3) = \binom{5}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = 0.0322$$

. ◀

EXAMPLE 3.5

Let the r.v $X \sim b(7, p)$ and that $\Pr(X = 3) = \Pr(X = 4)$. Find μ and σ^2, and then evaluate $\Pr(X = 3)$.

Solution. Since X has a binomial pdf, $f(x) = \Pr(X = x) = \binom{7}{x}p^x q^{7-x}$, $x = 0, 1, \ldots, 7$ and since we have $\Pr(X = 3) = \Pr(X = 4)$, then

$$f(3) = f(4) \Rightarrow \binom{7}{3}p^3q^4 = \binom{7}{4}p^4q^3 \Rightarrow p = q \Rightarrow p = 1 - p \Rightarrow p = \frac{1}{2}.$$

Therefore,

$$\mu = np = \frac{7}{2} \text{ and } \sigma^2 = npq = \frac{7}{4}.$$

Then,

$$\Pr(X = 3) = \Pr(x \leq 3) - \Pr(x \leq 2) = 0.5 - 0.2266 = 0.2734.$$

◀

3.1.2 The Geometric Distribution

The random variable with the geometric probability distribution is associated with an experiment that shares some of the characteristics of a binomial experiment. This experiment also involves identical and independent trials, each of which can result in one of two outcomes: success or failure. The probability of success is equal to p and is constant from trial to trial. However, instead of the number of successes that occur in n trials, the geometric random variable X is the number of the trial on which the first success occurs. Thus, the experiment consists of a series of trials that concludes with the first success. Consequently, the experiment could end with the first trial if a success is observed on the very first trial, or the experiment could go on indefinitely.

***Definition*:** A random variable X is said to have a geometric probability distribution if

$$f(x) = q^{x-1}p, \quad x = 1, 2, 3, \ldots, \ 0 \leq p \leq 1,$$

or

$$f(x) = q^{x}p, \quad x = 0, 1, 2, \ldots, \ 0 \leq p \leq 1,$$

We denote to X as $X \sim Geo(p)$.

To prove that $f(x)$ is a valid pdf,

1. $f(x) > 0, \ \forall x \in \mathcal{A} = \{x : x = 1, 2, \ldots\}$.
2. We need to evaluate $\sum_x f(x)$, then

$$\sum_{x=0}^{\infty} q^x p = p\left(1 + q + q^2 + q^3 + \ldots\right) = p\frac{1}{1-q} = p\frac{1}{p} = 1$$

The geometric probability distribution is often used to model distributions of lengths of waiting times.

EXAMPLE 3.6

Suppose that the probability of engine malfunction during any one-hour period is $p = 0.02$. Find the probability that a given engine will survive two hours.

Solution. Letting Y denote the number of one-hour intervals until the first malfunction, we have

$$\Pr(\text{survive two hours}) = \Pr(Y \geq 3) = \sum_{y=3}^{\infty} f(y).$$

Because $\sum_{y=1}^{\infty} f(y) = 1$,

$$\begin{aligned}\Pr(\text{survive two hours}) &= 1 - \sum_{y=1}^{2} f(y) \\ &= 1 - p - qp = 1 - 0.02 - (0.98)(0.02) = 0.9604.\end{aligned}$$

◀

***The Cumulative Distribution Function*:** The cdf of r.v X which has a Geometric pdf is:

$$F(x) = \Pr(X \leq x) = \sum_{\tau=0}^{x} f(\tau) = \sum_{\tau=0}^{x} q^{\tau}p = p(1 + q + q^2 + q^3 + \cdots + q^x)$$

$$= \begin{cases} 0, & x < 0 \\ 1 - q^{x+1}, & 0 \leq x < \infty \\ 1, & x = \infty \end{cases}$$

***The Moment Generating Function*:** We can obtain the mgf of r.v X by evaluating the expectation $E\left[e^{tX}\right]$, that is

$$M(t) = E\left[e^{tX}\right] = \sum_x e^{tx} f(x) = \sum_{x=0}^{\infty} e^{tx} q^x p = p\left[1 + (qe^t) + (qe^t)^2 + \dots\right]$$
$$= \frac{p}{1 - qe^t}, \quad t \neq \ln\left(\frac{1}{q}\right).$$

***Mean and Variance*:** If X is a r.v with a geometric distribution, then we define the mean μ and variance σ^2 as

$$\mu = E[X] = \frac{q}{p} \text{ and } \sigma^2 = Var(X) = \frac{1-p}{p^2}.$$

As it was mention previously, we can prove that in three different ways. This time, we will use the property of $\psi(t) = \ln(M(t))$. Hence $\psi(t) = \ln M(t) = \ln p - \ln(1 - qe^t)$, and the derivative of of ψ gives

$$\psi'(t) = \frac{qe^t}{1 - qe^t} \Rightarrow \psi'(0) = \frac{q}{1-q} = \frac{q}{p} = \mu$$

and,

$$\psi''(t) = \frac{(1 - qe^t)qe^t - (qe^t)(qe^t)}{(1 - qe^t)^2} \Rightarrow \psi''(0) = \frac{(1-q)(q) + q^2}{(1-q)^2} = \frac{q}{p^2} = \sigma^2$$

EXAMPLE 3.7

Let the r.v $X \sim Geo(\frac{2}{3})$. Answer the following:

1. Write down the pdf, cdf and mgf of X.
2. Find μ and σ^2.
3. Evaluate $\Pr(X \geq 3)$ by using both the pdf and cdf.
4. Evaluate $\Pr(X \geq 5 | X \geq 2)$.

Solution. 1. ▪ pdf: $f(x) = \left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^x, \ \ x = 0, 1, 2, \dots.$

▪ cdf:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - \left(\frac{1}{3}\right)^{x+1}, & 0 \leq x < \infty \\ 1, & x = \infty \end{cases}$$

▪ mgf: $M(t) = \dfrac{2/3}{1 - \frac{1}{3}e^t} = \dfrac{2}{3 - e^t}.$

2. $\mu = \dfrac{q}{p} = \dfrac{1/3}{2/3} = \dfrac{1}{2}$ and $\sigma^2 = \dfrac{q}{p^2} = \dfrac{1/3}{4/9} = \dfrac{3}{4}.$

3. ▪ Using pdf: $\Pr(X \geq 3) = 1 - \Pr(X < 3) = 1 - \Pr(X \leq 2)$, then

$$\Pr(X \geq 3) = 1 - \sum_{x=0}^{2} f(x) = 1 - [f(0) + f(1) + f(2)]1 - \frac{2}{3}\left[1 + \frac{1}{3} + \frac{1}{9}\right] = \frac{1}{27}.$$

- Using cdf: $\Pr(X \geq 3) = 1 - \Pr(X \leq 2) = 1 - F(2) = 1 - 1 + \left(\frac{1}{3}\right)^3 = \frac{1}{27}$.

4. According to the conditional probability,

$$\Pr(X \geq 5 | X \geq 2) = \frac{\Pr(X \geq 5 \cap X \geq 2)}{\Pr(X \geq 2)} = \frac{\Pr(X \geq 5)}{\Pr(X \geq 2)} = \frac{1 - \Pr(X \leq 4)}{1 - \Pr(X \leq 1)} = \frac{(1/3)^5}{(1/3)^2} = \frac{1}{27}.$$

◀

3.1.3 Negative Binomial Distribution

A random variable with a negative binomial distribution originates from a context much like the one that yields the geometric distribution. Again, we focus on independent and identical trials, each of which results in one of two outcomes: success or failure. The probability p of success stays the same from trial to trial. The geometric distribution handles the case where we are interested in the number of the trial on which the first success occurs. What if we are interested in knowing the number of the trial on which the second, third, or fourth success occurs? The distribution that applies to the random variable X equal to the number of the trial on which the rth success occurs ($r = 2, 3, 4,$ etc.) is the negative binomial distribution.

***Definition*:** A r.v X is said to have negative binomial distribution, denoted by $X \sim Nb(r, p)$, if X has the pdf $f(x)$, such that:

$$f(x) = \binom{x-1}{r-1} p^r q^{x-r}, \quad x = r, r+1, r+2, \ldots, \ 0 \leq p \leq 1.$$

Or

$$f(x) = \binom{x+r-1}{x} p^r q^x, \quad x = 0, 1, 2, \ldots, \ 0 \leq p \leq 1.$$

***Note*:** The Maclaurian's series expansion of:

$$\begin{aligned}(1-a)^{-r} =& 1 + ra + \frac{r(r+1)}{2!}a^2 + \frac{r(r+1)(r+2)}{3!}a^3 + \ldots \\ =& \binom{r-1}{0}a^0 + \binom{r}{1}a^1 + \binom{r+1}{2}a^2 + \binom{r+2}{3}a^3 + \ldots \\ =& \sum_{x=0}^{\infty} \binom{x+r-1}{x} a^x, \quad r = 1, 2, 3, \ldots\end{aligned}$$

In order to verify that $f(x)$ is a valid pdf, we note that

1. $f(x) > 0, \forall x \in \mathcal{A} = \{x : x = 0, 1, 2, \ldots\}$.
2. $\sum\limits_x f(x) = \sum\limits_{x=0}^{\infty} \binom{x+r-1}{x} p^r q^x = p^r \sum\limits_{x=0}^{\infty} \binom{x+r-1}{x} q^x = p^r (1-q)^{-r} = 1.$

EXAMPLE 3.8

A geological study indicates that an exploratory oil well drilled in a particular region should strike oil with probability 0.2. Find the probability that the third oil strike comes on the fifth well drilled.

Solution. Assuming independent drilling and probability 0.2 of striking oil with any one well, let X denote the number of the trial on which the third oil strike occurs. Then it is reasonable to assume that X has a negative binomial distribution with $p = 0.2$. Because we are interested in $r = 3$ and $x = 5$,

$$\Pr(X = 5) = f(5) = \binom{4}{2}(0.2)^3(0.8)^2 = 6(0.008)(0.64) = 0.0307.$$

◀

***The Moment Generating Function*:** The mgf of X could be presented as:

$$M(t) = E\left[e^{tX}\right] = \sum_{x=0}^{\infty} e^{tx}\binom{x+r-1}{x}p^r q^x = p^r \sum_{x=0}^{\infty}\binom{x+r-1}{x}(qe^t)^x = p^r(1-qe^t)^{-r}$$

$$M(t) = \left(\frac{p}{1-qe^t}\right)^r, \quad t \neq -\ln q.$$

***Mean and Variance*:** If X is a random variable with a negative binomial distribution,

$$\mu = E[X] = \frac{rq}{p} \text{ and } \sigma^2 = Var(X) = \frac{rq}{p^2}$$

EXAMPLE 3.9

A large stockpile of used pumps contains 20% that are in need of repair. A maintenance worker is sent to the stockpile with three repair kits. She selects pumps at random and tests them one at a time. If the pump works, she sets it aside for future use. However, if the pump does not work, she uses one of her repair kits on it. Suppose that it takes 10 minutes to test a pump that is in working condition and 30 minutes to test and repair a pump that does not work. Find the mean and variance of the total time it takes the maintenance worker to use her three repair kits.

Solution. Let X denote the number of the trial on which the third nonfunctioning pump is found. It follows that X has a negative binomial distribution with $p = 0.2$. Thus, $E(X) = 3/(0.2) = 15$ and $Var(X) = 3(0.8)/(0.2)2 = 60$. Because it takes an additional 20 minutes to repair each defective pump, the total time necessary to use the three kits is

$$T = 10X + 3(20).$$

Therefore, the expected time is,

$$E[T] = 10E[X] + 60 = 10(15) + 60 = 210,$$

and

$$Var(T) = 10^2 Var(X) = (100)(60) = 6000.$$

Thus, the total time necessary to use all three kits has mean 210 and standard deviation $\sqrt{6000} = 77.46$.

◀

3.1.4 The Hypergeometric Probability Distribution

Suppose that a population contains a finite number N of elements that possess one of two characteristics. Thus, r of the elements might be red and $b = N - r$, black. A sample of n elements is randomly selected from the population, and the random variable of interest is X, the number of red elements in the sample. This random variable has what is known as the hypergeometric probability distribution. For example, the number of workers who are women, X, in Example 2.2 has the hypergeometric distribution.

***Definition*:** A random variable X is said to have a hypergeometric probability distribution if and only if

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, 2, \ldots, n \text{ and } x \leq r,\ n - y \leq N - r.$$

1. We can easily notice that $f(x) \geq 0, \forall x$, since $\binom{a}{b} > 0, if\, a > b$, and $\binom{a}{b} = 0$, if $b > a$.

2. Notice that:

$$\sum_{i=0}^{n} \binom{r}{i}\binom{N-r}{n-i} = \binom{N}{n}.$$

Therefore,

$$\sum_{x=0}^{n} \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{\binom{N}{n}}{\binom{N}{n}} = 1.$$

This verifies that $f(x)$ is a valid pdf of the r.v X.

EXAMPLE 3.10

An important problem encountered by personnel directors and others faced with the selection of the best in a finite set of elements is exemplified by the following scenario. From a group of 20 Ph.D. engineers, 10 are randomly selected for employment. What is the probability that the 10 selected include all the 5 best engineers in the group of 20?

Solution. For this example $N = 20$, $n = 10$, and $r = 5$. That is, there are only 5 in the set of 5 best engineers, and we seek the probability that $Y = 5$, where Y denotes the number of best engineers among the ten selected. Then

$$f(y) = \frac{\binom{5}{5}\binom{15}{5}}{\binom{20}{10}} = \left(\frac{15!}{5!10!}\right)\left(\frac{10!10!}{20!}\right) = \frac{21}{1292} = 0.0162.$$

◀

***Mean and Variance*:** If X is a random variable with a hypergeometric distribution,

$$\mu = E[X] = \frac{nr}{N}, \text{ and } \sigma^2 = Var(X) = n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right).$$

EXAMPLE 3.11

An industrial product is shipped in lots of 20. Testing to determine whether an item is defective is costly, and hence the manufacturer samples his production rather than using a 100% inspection plan. A sampling plan, constructed to minimize the number of defectives shipped to customers, calls for sampling five items from each lot and rejecting the lot if more than one defective is observed. (If the lot is rejected, each item in it is later tested.) If a lot contains four defectives, what is the probability that it will be rejected? What is the expected number of defectives in the sample of size 5? What is the variance of the number of defectives in the sample of size 5?

Solution. Let X equal the number of defectives in the sample. Then $N = 20$, $r = 4$, and $n = 5$. The lot will be rejected if $X = 2, 3$, or 4. Then

$$\begin{aligned}\Pr(\text{rejecting the lot}) &= \Pr(X \geq 2) = f(2) + f(3) + f(4)\\ &= 1 - f(0) - f(1) = 1 - \frac{\binom{4}{0}\binom{16}{5}}{\binom{20}{5}} - \frac{\binom{4}{1}\binom{16}{4}}{\binom{20}{4}}\\ &= 1 - 0.2817 - 0.4694 = 0.2487\end{aligned}$$

The mean and variance of the number of defectives in the sample of size 5 are

$$\mu = \frac{(5)(4)}{20} = 1 \text{ and } \sigma^2 = 5\left(\frac{4}{20}\right)\left(\frac{20-4}{20}\right)\left(\frac{20-5}{20-1}\right) = 0.632.$$

◀

3.1.5 The Poisson Distribution

The Poisson probability distribution often provides a good model for the probability distribution of the number X of rare events that occur in space, time, volume, or any other dimension, where λ is the average value of X. Examples of random variables with approximate Poisson distributions are the number of telephone calls handled by a switchboard in a time interval, the number of radioactive particles that decay in a particular time period, the number of errors a typist makes in typing a page, and the number of automobiles using a freeway access ramp in a ten-minute interval.

***Definition*:** A random variable X is said to have a Poisson probability distribution if and only if

$$f(x) = \frac{\lambda^x}{x!}e^{-\lambda}, \quad x = 0, 1, 2, \ldots, \ \lambda > 0,$$

and is denoted by $X \sim P(\lambda)$

1. $f(x) > 0, \forall x$.
2. Since $e^{\lambda} = \sum\limits_{x=0}^{\infty} \frac{\lambda^x}{x!}$, then

$$\sum_x f(x) = \sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1.$$

EXAMPLE 3.12

Suppose that a random system of police patrol is devised so that a patrol officer may visit a given beat location $X = 0, 1, 2, 3, \ldots$ times per half-hour period, with each location being visited an average of once per time period. Assume that X possesses, approximately, a Poisson probability distribution. Calculate the probability that the patrol officer will miss a given location during a half-hour period. What is the probability that it will be visited once? Twice? At least once?

Solution. For this example the time period is a half-hour, and the mean number of visits per half-hour interval is $\lambda = 1$. Then

$$f(x) = \frac{(1)^x e^{-1}}{x!} = \frac{e^{-1}}{x!}, \quad x = 0, 1, 2, \ldots.$$

The event that a given location is missed in a half-hour period corresponds to $(X = 0)$, and

$$\Pr(X = 0) = f(0) = \frac{e^{-1}}{0!} = e^{-1} = 0.368.$$

Similarly,

$$\Pr(X = 1) = f(1) = \frac{e^{-1}}{1!} = e^{-1} = 0.368.$$

and

$$\Pr(X = 2) = f(2) = \frac{e^{-1}}{2!} = \frac{e^{-1}}{2} = 0.184.$$

The probability that the location is visited at least once is the event $(X \geq 1)$. Then

$$\Pr(X \geq 1) = \sum_{x=1}^{\infty} f(x) = 1 - f(0) = 1 - e^{-1} = 0.632.$$

◀

***The Cumulative Distribution Function*:** The case in Poisson distribution is similar to the Binomial distribution. The cdf of X is very complicated, as

$$F(x) = \Pr(X \leq x) = \sum_{\tau=0}^{x} \frac{\lambda^{\tau}}{\tau!} e^{-\lambda}.$$

The approximated value of Piosson cdf could be found in tables (2).

***The Moment Generating Function*:** The mgf of r.v X is:

$$M(t) = E\left[e^{tX}\right] = \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x}{x!} e^{-\lambda} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t}$$

$$M(t) = e^{\lambda(e^t - 1)}$$

***Mean and Variance*:** If X is a r.v possessing a Poisson distribution with parameter λ, then

$$\mu = E[X] = \lambda, \ \ and \ \ \sigma^2 = Var(X) = \lambda.$$

In order to prove that, set $\psi(t) = \ln M(t) = \lambda(e^t - 1)$. Then

$$\psi'(t) = \lambda e^t \Rightarrow \psi'(0) = \lambda e^0 = \lambda = \mu$$
$$\psi''(t) = \lambda e^t \Rightarrow \psi''(0) = \lambda e^0 = \lambda = \sigma^2$$

EXAMPLE 3.13

Let the r.v $X \sim P(2)$.Find μ and σ^2. Evaluate $\Pr(X \geq 2)$, $\Pr(2 \leq X \leq 5)$ and $\Pr(X = 3)$.

Solution. Since $\lambda = 2$, then: $\mu = \sigma^2 = 2$.

- $\Pr(X \geq 2) = 1 - \Pr(X \leq 1) = 1 - \sum_{x=0}^{1} f(x) = 1 - (f(0) + f(1)) = 1 - 0.406 = 0.594.$
- $\Pr(2 \leq X \leq 5) = F(5) - F(1) = 0.9834 - 0.4060 = 0.5774.$
- $\Pr(X = 3) = f(3) = e^{-2}\frac{2^3}{3!} = e^{-2}\frac{4}{3} = 0.1804.$

◀

EXAMPLE 3.14

Industrial accidents occur according to a Poisson process with an average of three accidents per month. During the last two months, ten accidents occurred. Does this number seem highly improbable if the mean number of accidents per month, μ, is still equal to 3?

Solution. The number of accidents in two months, X, has a Poisson probability distribution with mean $\lambda^* = 2(3) = 6$. The probability that X is as large as 10 is

$$\Pr(X \geq 10) = \sum_{x=10}^{\infty} \frac{6^x}{x!} e^{-6}.$$

The tedious calculation required to find $\Pr(X \geq 10)$ can be avoided by using Table 2,

$$\Pr(X \geq 10) = 1 - \Pr(X \leq 9) = 1 - 0.9161 = 0.0839$$

It is not highly improbable to happen. ◄

***Theorem*:** Let the r.v $X \sim b(n,p)$. For large $n \to \infty$ and small $p \to 0$. Then X could be approximated as $\lim_{n\to\infty} b(n,p) = P(\lambda)$, where $\lambda = np$.

EXAMPLE 3.15

Let the r.v $X \sim b(3000, 0.001)$. Find $\Pr(X = 5)$.

Solution. Since $X \sim b(n,p)$ with a large value of $n = 3000$ and small value of $p = 0.001$. Then according to the above theorem, we can approximate that $X \sim \lambda$, where $\lambda = np = (3000)(0.001) = 3$. Then:

$$\Pr(X = 5) = \Pr(X \leq 5) - \Pr(X \leq 4) = 0.9161 - 0.8153 = 0.1008.$$

◄

3.2 Continuous Distribution

In this section, some of the most important and popular distributions for continuous r.v are presented. Both the pdf and cdf is derived and some important properties and mathematical expectation for these distribution are obtained.

3.2.1 The Uniform Distribution

Suppose that a bus always arrives at a particular stop between 8:00 and 8:10 A.M. and that the probability that the bus will arrive in any given subinterval of time is proportional only to the length of the subinterval. That is, the bus is as likely to arrive between 8:00 and 8:02 as it is to arrive between 8:06 and 8:08. Let X denote the length of time a person must wait for the bus if that person arrived at the bus stop at exactly 8:00. If we carefully measured in minutes how long after 8:00 the bus arrived for several mornings, we could develop a relative frequency histogram for the data. From the description just given, it should be clear that the relative frequency with which we observed a value of X between 0 and 2 would be approximately the same as the relative frequency with which we observed a value of X between 6 and 8. The random variable X just discussed is an example of a random variable that has a uniform distribution. The general form for the density function of a random variable with a uniform distribution is as follows.

***Definition*:** If $a < b$, a r.v X is said to have a continuous uniform pdf on the interval (a,b) if and only if the density function of X is:

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b$$

If X has a uniform distribution, then X is denoted by $X \sim U(a,b)$.

In order to check that $f(x)$ is a pdf, we note that:

1. $f(x) > 0, \forall x \in (a,b)$.
2. $\int_x f(x)dx = \int_a^b \frac{dx}{b-a} = \frac{x}{b-a}\Big|_a^b = \frac{b-a}{b-a} = 1$.

EXAMPLE 3.16

Arrivals of customers at a checkout counter follow a Poisson distribution. It is known that, during a given 30-minute period, one customer arrived at the counter. Find the probability that the customer arrived during the last 5 minutes of the 30-minute period.

Solution. As just mentioned, the actual time of arrival follows a uniform distribution over the interval of $(0, 30)$. If X denotes the arrival time, then

$$\Pr(25 < X < 30) = \int_{25}^{30} \frac{1}{30} dx = \frac{30-25}{30} = \frac{5}{30} = \frac{1}{6}.$$

The probability of the arrival occurring in any other 5-minute interval is also 1/6. ◀

***The Cumulative Distribution Function*:** The cdf of X is derived as

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f(\tau) d\tau = \int_{a}^{x} \frac{d\tau}{b-a} = \frac{\tau}{b-a}\Big|_a^x$$

Therefore:

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & x \geq b \end{cases}$$

***Mean and Variance*:** If $a < b$ and X is a r.v uniformly distributed on the interval (a, b), then

$$\mu = E[X] = \frac{a+b}{2} \text{ and } \sigma^2 = Var(X) = \frac{(b-a)^2}{12}.$$

It is easy to evaluate the mean and variance of X by direct definition as

$$E[X] = \int_a^b \frac{x}{b-a} dx = \frac{x^2}{2(b-a)}\Big|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}.$$

also

$$E[X^2] = \int_a^b \frac{x^2}{b-a} dx = \frac{x^3}{3(b-a)}\Big|_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ab+a^2}{3},$$

therefore

$$\sigma^2 = E[X^2] - \mu^2 = \frac{b^2+ab+a^2}{3} - \left(\frac{b+a}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

EXAMPLE 3.17

Let the r.v $X \sim U(a, b)$ with $\mu = 1$ and $\sigma^2 = \frac{4}{3}$. Find $\Pr(X < 0)$.

Solution. Since the mean $\mu = 1$ and

$$\mu = \frac{a+b}{2} = 1 \Rightarrow a + b = 2 \tag{3.1}$$

also, the variance $\sigma^2 = \frac{4}{3}$, and

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{4}{3} \Rightarrow (b-a)^2 = 16 \Rightarrow b - a = \pm 4 \tag{3.2}$$

Consider $a+b=2$ and $b-a=-4$. Hence $b=-1$ and $a=3$. This is non applicable because $X \sim U(a,b)$ and therefore $a<b$. Then, $a+b=2$ and $b-a=4$. Hence, $b=3$ and $a=-1$.

$$\therefore f(x)=\frac{1}{b-a}=\frac{1}{3-(-1)}=\frac{1}{4}, \quad -1<x<3.$$

Therefore,

$$\Pr(x<0)=\int_{-1}^{0} f(x)dx=\frac{1}{4}\int_{-1}^{0} dx=\frac{1}{4}.$$

◀

EXAMPLE 3.18

If a parachutist lands at a random point on a line between markers A and B, find the probability that he is closer to A than to B. Find the probability that his distance to A is more than three times his distance to B.

Solution. Consider the landing point X is a r.v and $X \sim (A,B)$, then for the first part of the question, we assume that $X-A<B-X$, and

$$\Pr(X-A<B-X)=\Pr(2X<B+A)=\Pr\left(X<\frac{B+A}{2}\right)=F\left(\frac{B+A}{2}\right)$$
$$=F\left(\frac{B+A}{2}\right)=\frac{\frac{B+A}{2}-A}{B-A}=\frac{\frac{A+B-2A}{2}}{B-A}=\frac{1}{2}.$$

For the second part, assuming that $X-A>3(B-X)$, and that

$$\Pr(X-A>3(B-X))=\Pr(4X>3B+A)=\Pr\left(X>\frac{3B+A}{4}\right)=1-F\left(\frac{3B+A}{4}\right)$$
$$=1-F\left(\frac{3B+A}{4}\right)=1-\frac{\frac{3B+A}{4}-A}{B-A}=1-\frac{\frac{3B+A-4A}{4}}{B-A}=1-\frac{\frac{3(B-A)}{4}}{B-A}=1-\frac{3}{4}=\frac{1}{4}.$$

◀

3.2.2 Gamma Distribution

Some random variables are always non-negative and for various reasons yield distributions of data that are skewed (non-symmetric) to the right. That is, most of the area under the density function is located near the origin, and the density function drops gradually as x increases. A skewed probability density function is shown in Figure 3.1

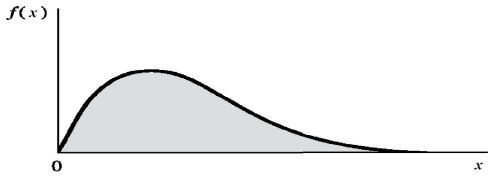

Figure 3.1 A skewed probability density function

A family of pdf's that yields a wide variety of skewed distributional shapes is the gamma family. To define the family of gamma distributions, we first need to introduce a function that plays an important role in many branches of mathematics.

***Gamma Function*:** For any $\alpha > 0$, the gamma function $\Gamma(\alpha)$ is defined by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

The most important properties of the gamma function are the following:

1. For any $\alpha > 0$, $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$. (via integration by parts).
2. For any positive integer, n, $\Gamma(n) = (n-1)!$.
3. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

***Probability Density Function*:** A random variable X is said to have a gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ if and only if the density function of X is

$$f(x) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^\alpha \Gamma(\alpha)}, \quad 0 \le x < \infty$$

and X is denoted by $X \sim G(\alpha, \beta)$. To prove that $f(x)$ is a valid pdf, we need to check:

1. $f(x) > 0, \forall x \in [0, \infty)$.
2. to prove the unity of integration, we use the gamma function, then

$$\int_x f(x)dx = \int_0^\infty \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} dx$$

Let $y = \frac{x}{\beta} \Rightarrow x = \beta y \Rightarrow dx = \beta dy$. Then

$$\int_0^\infty f(x)dx = \int_0^\infty \frac{1}{\beta^\alpha \Gamma(\alpha)} (\beta y)^{\alpha-1} e^{-y} \beta dy = \frac{1}{\Gamma(\alpha)} \int_0^\infty y^{\alpha-1} e^{-y} dy = \frac{1}{\Gamma(\alpha)} \Gamma(\alpha) = 1.$$

Therefore, $f(x)$ is a pdf.

***Cumulative Distribution Function*:** In the special case when α is an integer, cdf of a gamma distributed r.v can be expressed as a sum of certain Poisson probabilities. However, if α is not an integer, it is impossible to give a closed-form expression for

$$\int_0^x \frac{1}{\beta^\alpha \Gamma(\alpha)} \tau^{\alpha-1} e^{-\frac{\tau}{\beta}} d\tau$$

This integral is called an incomplete gamma function, and except when $\alpha = 1$ (an exponential distribution), it is impossible to obtain areas under the gamma density function by direct integration. There are not many tables of $F(x)$ available; in Table A.4, we present a small tabulation for $\alpha = 1, 2, \ldots, 10$, $\beta = 1$ and $x = 1, 2, \ldots, 15$.

***Moment Generating Function*:** The moment generating function of a gamma random variable is

$$M(t) = \frac{1}{(1-\beta t)^\alpha}, \quad t < \frac{1}{\beta}.$$

We use the definition of the mgf to prove it, such as

$$M(t) = E[e^{tX}] = \int_0^\infty e^{tx} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-\frac{1}{\beta}(1-\beta t)x} dx$$

Set $y = \left(\frac{1-\beta t}{\beta}\right) x \Rightarrow x = \left(\frac{\beta}{1-\beta t}\right) y \Rightarrow dx = \left(\frac{\beta}{1-\beta t}\right) dy$, then

$$M(t) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty \left(\frac{\beta y}{1-\beta t}\right)^{\alpha-1} e^{-y} \left(\frac{\beta}{1-\beta t}\right) dy = \frac{1}{\beta^\alpha \Gamma(\alpha)} \left(\frac{\beta}{1-\beta t}\right)^\alpha \int_0^\infty y^{\alpha-1} e^{-y} dy = \frac{1}{(1-\beta t)^\alpha}.$$

***Mean and Variance*:** If X has a gamma distribution with parameters α and β, then

$$\mu = E[X] = \alpha\beta, \text{ and } \sigma^2 = Var(X) = \alpha\beta^2.$$

We can prove that either with direct integration or using the logarithm for the mgf.

$$\psi(t) = \ln(M(t)) = -\alpha \ln(1-\beta t) \Rightarrow \psi'(t) = \frac{\alpha\beta}{(1-\beta t)} \Rightarrow \psi'(0) = \alpha\beta = \mu.$$

and

$$\psi''(t) = \frac{\alpha\beta^2}{(1-\beta t)^2} \Rightarrow \psi''(0) = \alpha\beta^2 = \sigma^2.$$

EXAMPLE 3.19

Let the r.v $X \sim G(3, 2)$. Write down the pdf and mgf of X. Find the mean μ and variance σ^2. Evaluate the probability $\Pr(12.6 < x < 16.8)$.

Solution. Since we have $\alpha = 3$ and $\beta = 2$, then the pdf of X is

$$f(x) = \frac{1}{2^3 \Gamma(3)} x^{3-1} e^{-\frac{x}{2}} = \frac{1}{16} x^2 e^{-\frac{x}{2}}, 0 < x < \infty.$$

then, the mgf is

$$M(t) = \frac{1}{(1-2t)^3}.$$

also, $\mu = \alpha\beta = 6$ and $\sigma^2 = \alpha\beta^2 = 12$. Finally,

$$\Pr(12.6 < X < 16.8) = \int_{12.6}^{16.8} f(x) = \frac{1}{16} \int_{12.6}^{16.8} x^2 e^{-\frac{x}{2}} dx = 0.04.$$

◀

EXAMPLE 3.20

Find the constant c so that the function $f(x) = cx^5e^{-3x}$, $0 < x < \infty$ is a pdf.

Solution. Since $f(x)$ is a pdf, then

$$1 = \int_0^\infty f(x)dx = c\int_0^\infty x^5e^{-3x}dx = c\frac{\Gamma(6)}{3^6}\int_0^\infty \frac{3^6}{\Gamma(6)}x^{6-1}e^{-3x}dx = \frac{5!c}{3^6}$$

Therefore, $c = \frac{3^6}{5!}$. ◄

EXAMPLE 3.21

Suppose the survival time X in weeks of a randomly selected male mouse exposed to 240 rads of gamma radiation has a gamma distribution with $\alpha = 8$ and $\beta = 1$. What is the expected survival time for the mouse? what is the probability that the mouse could survive between 6 and 12 weeks and the probability of it survives at least 3 weeks?

Solution. Since $X \sim G(8, 1)$, then the expected survival time in weeks is $E[X] = \alpha\beta = (8)(1) = 8$ weeks.
The probability that a mouse survives between 6 and 12 weeks is

$$\Pr(6 \leq X \leq 12) = \Pr(X \leq 12) - \Pr(X \leq 6) = 0.911 - 0.256 = 0.655.$$

The probability that a mouse survives at least 3 weeks is

$$\Pr(X \geq 3) = 1 - \Pr(X \leq 3) = 1 - 0.012 = 0.998.$$

◄

3.2.3 The Exponential Distribution

The family of exponential distributions provides probability models that are widely used in engineering and science disciplines. For example, the exponential density function is often useful for modelling the length of life of electronic components.

***Definition*:** A r.v X is said to have an exponential distribution with parameter $\lambda > 0$ if the pdf of X is

$$f(x) = \frac{1}{\lambda}e^{-\frac{x}{\lambda}}, \quad x \geq 0.$$

The r.v is denoted by $X \sim Exp(\lambda)$. In order to prove that $f(x)$ is a valid pdf, we can simply prove

1. $f(x) > 0, \forall x \in (0, \infty)$.
2. It is easy to prove that the integral

$$\int_x f(x)dx = \int_0^\infty \frac{1}{\lambda}e^{-\frac{x}{\lambda}}dx = -e^{-\frac{x}{\lambda}}\Big|_0^\infty = -(e^{-\infty} - e^0) = 1.$$

***Note*:** The exponential pdf is a special case of the general gamma pdf in which $\alpha = 1$ and β has been replaced by $1/\lambda$.

***The Cumulative Distribution Function*:** The cdf of a r.v that is exponentially distributed could be evaluate as

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^x f(\tau)d\tau = \int_0^x \frac{1}{\lambda}e^{-\frac{\tau}{\lambda}}d\tau = -e^{-\frac{\tau}{\lambda}}\Big|_0^x = 1 - e^{-\frac{x}{\lambda}},$$

therefore

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-\frac{x}{\lambda}}, & 0 < x < \infty \\ 1, & x = \infty \end{cases}$$

***The Moment Generating Function*:** The mgf of X is also evaluated for exponential distributed r.v as

$$M(t) = E[e^{tX}] = \int_x e^{tx} f(x)dx = \int_0^\infty e^{tx} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx = \int_0^\infty \frac{1}{\lambda} e^{-\frac{1}{\lambda}(1-\lambda t)x} = -\frac{e^{-\frac{1}{\lambda}(1-\lambda t)x}}{(1-\lambda t)}\Big|_0^\infty$$
$$= \frac{1}{1-\lambda t}, \quad t < \frac{1}{\lambda}.$$

***Mean and Variance*:**If X is an exponential r.v with parameter λ, then

$$\mu = E[X] = \lambda \text{ and } \sigma^2 = Var(X) = \lambda^2.$$

The proof simply follows directly from the mean and variance of the gamma distribution with $\alpha = 1$.

***Note*:** When $\lambda = 1$, then r.v $X \sim Exp(1)$ is said to have a standard exponential distribution with pdf $f(x) = e^{-x},\ 0 < x < \infty$.

EXAMPLE 3.22

Let the r.v $X \sim Exp(\lambda)$. If $\Pr(X \leq 1) = \Pr(X > 1)$. Find μ and σ^2.

Solution. Since $\Pr(X \leq 1) = \Pr(x > 1) = 1 - \Pr(X \leq 1) \Rightarrow 2\Pr(X \leq 1) = 1 \Rightarrow \Pr(X \leq 1) = \frac{1}{2}$. Therefore, using the cdf $F(x) = 1 - e^{-\frac{x}{\lambda}},\ 0 < x < \infty$, then

$$\Pr(X \leq 1) = F(1) = 1 - e^{-\frac{1}{\lambda}} = \frac{1}{2} \Rightarrow e^{-\frac{1}{\lambda}} = \frac{1}{2} \Rightarrow -\frac{1}{\lambda} = \ln\left(\frac{1}{2}\right) = \ln(1) - \ln(2) = 0 - \ln(2) = -\ln(2)$$
$$\Rightarrow \lambda = \frac{1}{\ln(2)}.$$

Hence, $\mu = \lambda = \frac{1}{\ln(2)}$ and $\sigma^2 = \lambda^2 = \left(\frac{1}{\ln(2)}\right)^2$. ◀

EXAMPLE 3.23

If a r.v $X \sim Exp(2)$. Find $\Pr(X \leq 1 | X \leq 2)$.

Solution.

$$\Pr(X \leq 1 | X \leq 2) = \frac{\Pr(X \leq 1 \cap X \leq 2)}{\Pr(X \leq 2)} = \frac{\Pr(X \leq 1)}{\Pr(X \leq 2)} = \frac{F(1)}{F(2)} = \frac{1 - e^{-\frac{1}{2}}}{1 - e^{-1}}.$$

◀

EXAMPLE 3.24

The response time X at an on-line computer terminal (the elapsed time between the end of a user's inquiry and the beginning of the system's response to that inquiry) has an exponential distribution with expected response time equal to 5 s. Then $E[X] = \lambda = 5$ s. Find the probability that the response time is at most 10 s and the probability that response time is between 5 and 10 s.

Solution. The probability that the response time is at most 10 s is

$$\Pr(X \leq 10) = F(10) = 1 - e^{-\frac{10}{5}} = 1 - e^{-2} = 0.865.$$

The probability that response time is between 5 and 10 s is

$$\Pr(5 \leq X \leq 10) = F(10) - F(5) = (1 - e^{-2}) - (1 - e^{-1}) = 0.233.$$

◀

EXAMPLE 3.25

One-hour carbon monoxide concentrations in air samples from a large city have an approximately exponential distribution with mean 3.6 ppm (parts per million).

a Find the probability that the carbon monoxide concentration exceeds 9 ppm during a randomly selected one-hour period.

b A traffic-control strategy reduced the mean to 2.5 ppm. Now find the probability that the concentration exceeds 9 ppm.

Solution. Since $X \sim Exp(\lambda)$, and the mean of carbon monoxide concentrations in air is 3.6, then

$$\Pr(X \geq 9) = 1 - \Pr(X \leq 9) = 1 - F(9) = 1 - (1 - e^{-\frac{9}{3.6}}) = 0.082$$

If the mean of carbon monoxide concentrations in air is reduced to 2.5, then

$$\Pr(X \geq 9) = 1 - \Pr(X \leq 9) = 1 - F(9) = 1 - (1 - e^{-\frac{9}{2.5}}) = 0.027.$$

◀

3.2.4 Chi-Square Distribution

Let ν be a positive integer. A random variable X is said to have a chi-square distribution with ν degrees of freedom (dof) if X is a gamma-distributed r.v with parameters $\alpha = \nu/2$ and $\beta = 2$. Such random variables occur often in statistical theory. The motivation behind calling the parameter ν the degrees of freedom of the Chi-square distribution rests on one of the major ways for generating a random variable with this distribution.

***The Probability Density Function*:** The r.v X is said to have a Chi-square distribution if it has a pdf

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

and is denoted by $X \sim \chi^2(\nu)$.

***The Cumulative Distribution Function*:** As the case of the gamma distribution, the cdf of chi-square distribution is indeterminable as the integral

$$F(x) = \Pr(X \leq x) = \int_0^x \frac{1}{2^{(\nu/2)}\Gamma(\nu/2)} \tau^{(\nu/2)-1} e^{-(\tau/2)} d\tau$$

is almost impossible to solve. However, values of this integral corresponding to specific values of x and ν are given in the tables.

***The Moment Generating Function*:** The mgf of the r.v $X \sim \chi^2(\nu)$ is

$$M(t) = E[e^{tX}] = \frac{1}{(1-2t)^{\nu/2}}, \quad t < \frac{1}{2}.$$

***Mean and Variance*:** Since $X \sim \chi^2(\nu)$ is equal to $X \sim G(\frac{\nu}{2}, 2)$, then the mean and variance are

$$\mu = E[X] = \nu \text{ and } \sigma^2 = Var(X) = 2\nu.$$

EXAMPLE 3.26

Let the r.v X has an mgf $M(t) = (1-2t)^{-8}$. What is the distribution of X? Find μ and σ^2. Evaluate $\Pr(6.97 < X < 26.3)$.

Solution. The given mgf is for a r.v $X \sim G(8, 2)$ or $X \sim \chi^2(16)$, with dof $\nu = 16$. Therefore $\mu = \nu = 16$ and $\sigma^2 = 2\nu = 32$.

The probability of

$$\Pr(7.96 < X < 26.3) = \Pr(X < 26.3) - \Pr(X < 7.96) = 0.95 - 0.05 = 0.9.$$

◀

3.2.5 The Beta Distribution

The beta density function is a two-parameter density function defined over the closed interval $0 \leq x \leq 1$. It is often used as a model for proportions, such as the proportion of impurities in a chemical product or the proportion of time that a machine is under repair.

***The Probability Density Function*:** A r.v X is said to have a beta probability distribution with parameters $\alpha > 0$ and $\beta > 0$, denoted by $X \sim Be(\alpha, \beta)$, if the density function of X is

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad 0 \leq x \leq 1,$$

where $B(\alpha, \beta)$, the beta function, is defined as

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

To prove that $f(x)$ is a pdf, we need to check that

1. $f(x) > r, \forall x \in (0, 1)$.

2. To show that $\int\limits_x f(x)dx = 1$. Consider the integral $I = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$. Successive integration by parts gives

$$\begin{array}{rcl} (1-x)^{\beta-1} & & x^{\alpha-1} \\ -(\beta-1)(1-x)^{\beta-2} & \searrow & \frac{x^{\alpha}}{\alpha} \\ (\beta-1)(\beta-2)(1-x)^{\beta-3} & \searrow & \frac{x^{\alpha+1}}{\alpha(\alpha+1)} \\ \vdots & \vdots & \frac{x^{\alpha+2}}{\alpha(\alpha+1)(\alpha+2)} \\ \pm(\beta-1)(\beta-2)\ldots[\beta-(\beta-1)](1-x)^{\beta-\beta} & \searrow & \vdots \\ 0 & & \frac{x^{\alpha+\beta-1}}{\alpha(\alpha+1)\ldots(\alpha+\beta-1)} \end{array}$$

$$\begin{aligned} \therefore I &= \frac{1}{\alpha}x^{\alpha}(1-x)^{\beta-1} + \frac{(\beta-1)}{\alpha(\alpha+1)}x^{\alpha+1}(1-x)^{\beta-2} + \frac{(\beta-1)(\beta-2)}{\alpha(\alpha+1)(\alpha+2)}x^{\alpha+2}(1-x)^{\beta-3} + \ldots \\ &\quad + \frac{(\beta-1)(\beta-2)\ldots 3\times 2\times 1}{\alpha(\alpha+1)(\alpha+2)\ldots(\alpha+\beta-1)}x^{\alpha+\beta-1}\Big|_0^1 \\ &= \frac{(\beta-1)(\beta-2)\ldots 3\times 2\times 1}{\alpha(\alpha+1)(\alpha+2)\ldots(\alpha+\beta-1)} = \frac{(\beta-1)!}{\alpha(\alpha+1)(\alpha+2)\ldots(\alpha+\beta-1)} \times \frac{(\alpha-1)(\alpha-2)\ldots 3\times 2\times 1}{(\alpha-1)(\alpha-2)\ldots 3\times 2\times 1} \\ \therefore I &= \frac{(\beta-1)!(\alpha-1)!}{(\alpha+\beta-1)!} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}. \end{aligned}$$

Thus, the integral of $f(x)$ is

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = 1$$

***Note*:** In the case of $\alpha = \beta = 1$, the beta distribution reduces to be a standard uniform, that is $Be(1,1) \equiv U(0,1)$.

***The Cumulative Distribution Function*:** The cdf for the beta r.v is commonly called the incomplete beta function and is denoted by

$$F(x) = \int_0^x \frac{\tau^{\alpha-1}(1-\tau)^{\beta-1}}{B(\alpha,\beta)}d\tau.$$

A tabulation of $F(x)$ is given in Tables of the Incomplete Beta Function (Pearson, 1968). When α and β are both positive integers, $F(x)$ is related to the binomial probability function.

***Mean and Variance*:** If X is a beta distributed r.v with parameters $\alpha > 0$ and $\beta > 0$, then

$$\mu = E[X] = \frac{\alpha}{\alpha+\beta} \text{ and } \sigma^2 = Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

We can derive the rth moment by the direct definition and by the use of the same technique we used to prove the unity of the pdf integration, then we can find that

$$E[X^r] = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+r)}{\Gamma(\alpha)\Gamma(\alpha+\beta+r)}, \quad r = 1,2,3,\ldots$$

In order to derive the mean and variance, set $r = 1, 2$ to get

$$r = 1 \Rightarrow \mu = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+1)}{\Gamma(\alpha)\Gamma(\alpha+\beta+1)} = \frac{(\alpha+\beta-1)!\,\alpha!}{(\alpha-1)!\,(\alpha+\beta)!} = \frac{\alpha}{\alpha+\beta}.$$

$$r = 2 \Rightarrow \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)} \Rightarrow \sigma^2 = E[X^2] - \mu^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}.$$

EXAMPLE 3.27

Find the constant c in the pdf: $f(x) = c\,x^{19}(1-x)^{29},\ \ 0 < x < 1$.

Solution. Since $f(x)$ is a pdf, then

$$1 = c\int_0^1 x^{19}(1-x)^{29}dx = c\frac{\Gamma(20)\Gamma(30)}{\Gamma(50)}\int_0^1 \frac{\Gamma(50)}{\Gamma(20)\Gamma(30)}x^{20-1}(1-x)^{30-1}dx$$
$$\Rightarrow c = \frac{\Gamma(50)}{\Gamma(20)\Gamma(30)}.$$

◀

EXAMPLE 3.28

A gasoline wholesale distributor has bulk storage tanks that hold fixed supplies and are filled every Monday. Of interest to the wholesaler is the proportion of this supply that is sold during the week. Over many weeks of observation, the distributor found that this proportion could be modelled by a beta distribution with $\alpha = 4$ and $\beta = 2$. Find the probability that the wholesaler will sell at least 90% of his stock in a given week.

Solution. If Y denotes the proportion sold during the week, then

$$f(y) = \frac{\Gamma(4+2)}{\Gamma(4)\Gamma(2)}y^3(1-y), \quad 0 \le y \le 1.$$

and therefore,

$$\Pr(Y > 0.9) = \int_{0.9}^1 f(y)dy = 20\int_{0.9}^1 y^3(1-y)dy = 20\left\{\frac{y^4}{4}\Big|_{0.9}^1 - \frac{y^5}{5}\Big|_{0.9}^1\right\} = (20)(0.004) = 0.08.$$

It is not very likely that 90% of the stock will be sold in a given week. ◀

3.2.6 The Normal Distribution

The normal distribution with the familiar bell shape is the most important one in all of probability and statistics. Many numerical populations have distributions that can be fit very closely by an appropriate normal curve. Examples include heights, weights, and other physical characteristics, measurement errors in scientific experiments, measurements on fossils, reaction times in psychological experiments, measurements of intelligence and aptitude, scores on various tests, and numerous economic measures and indicators. Even when the underlying distribution is discrete, the normal curve often gives an excellent approximation. In addition, even when individual variables themselves are not normally distributed, sums and averages of the variables will under suitable conditions have approximately a normal distribution; this is the content of the Central Limit Theorem discussed in the following chapters.

***Probability Density Function*:** A continuous r.v X is said to have a normal distribution with parameters μ and σ^2, where $-\infty < \mu < \infty$ and $\sigma > 0$, if the pdf of X is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad \infty < x < \infty.$$

The r.v X is denoted by $X \sim N(\mu, \sigma^2)$. Here is a proof that the normal curve satisfies the pdf requirement:

1. $f(x) > 0, \forall x \in (-\infty, \infty)$.

2. In order to satisfy $\int_x f(x)dx = 1$, we will need to use some special technique. First of all consider the integral $I = \int_{-\infty}^{\infty} e^{-x^2/2}dx$. Then

$$I^2 = \left\{\int_{-\infty}^{\infty} e^{-x^2/2}dx\right\}\left\{\int_{-\infty}^{\infty} e^{-y^2/2}dy\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}dxdy$$

Use the Polar transformation by setting, $x = r\cos\theta,\ 0 \le r < \infty$ and $y = r\sin\theta,\ 0 \le \theta \le 2\pi$. Also, the Jacobian for the transformation is

$$J = \frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r.$$

Then, $dxdy = |J|drd\theta = rdrd\theta$. This implies that the integral I^2 can be written as

$$I^2 = \int_{r=0}^{\infty}\int_{\theta=0}^{2\pi} e^{-r^2/2}rdrd\theta = 2\pi\int_{r=0}^{\infty} re^{-r^2/2}dr = -(2\pi)\ \left[e^{-r^2/2}\right]_0^{\infty} = 2\pi$$

Hence, $I = \int_{-\infty}^{\infty} e^{-x^2/2}dx = \sqrt{2\pi}$. Back to the Normal distribution function,

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx,$$

set $y = \frac{x-\mu}{\sigma} \Rightarrow x = \sigma y + \mu$, then $\Rightarrow dx = \sigma dy$ and

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-y^2/2}\sigma dy = \frac{1}{\sqrt{2\pi}}\sqrt{2\pi} = 1.$$

This proves that $f(x)$ is a pdf.

***Cumulative Distribution Function*:** There is no a closed form expression for the cdf of a normal distributed r.v because the integral $F(x) = \Pr(X \le x) = \int_{-\infty}^{x} f(\tau)d\tau$ has no analytic solution. hence, its evaluation requires the use of numerical integration techniques. However, in literature, tables are approximated for the case of standard Normal distribution, that is $X \sim N(0,1)$.

***Definition*:** The normal distribution with mean $\mu = 0$ and $\sigma^2 = 1$ is called the standard normal distribution. A r.v that has a standard normal distribution is called a standard normal random variable and will be denoted by Z. The pdf of Z is

$$f(z) = \frac{1}{\sqrt{2\pi}}\ e^{-z^2/2}, \quad -\infty < z < \infty.$$

The cdf of Z is $F(z) = \Pr(Z \le z) = \int_{-\infty}^{z} f(\tau)d\tau$, which we will denote by $\Phi(z)$.

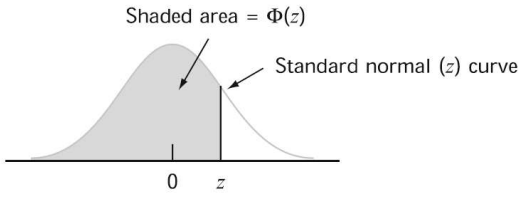

Figure 3.2 Standard Normal Distribution

EXAMPLE 3.29

Let Z denote a normal random variable with mean 0 and standard deviation 1.

1. Find $\Pr(Z > 1.25)$.
2. Find $\Pr(-0.38 \leq Z \leq 1.25)$.

Solution.

1. $\Pr(Z > 1.25) = 1 - \Pr(Z \leq 1.25) = 1 - \Phi(1.25)$, the probability $\Phi(z)$ can be found in tables and is $\Phi(1.25) = 0.8944$. Therefore $\Pr(Z > 1.25) = 1 - 0.8944 = 0.1056$.
2. $\Pr(-0.38 \leq Z \leq 1.25) = \Phi(1.25) - \Phi(-0.38) = 0.8944 - 0.3520 = 0.5424$.

◀

***Nonstandard Normal Distributions*:** When $X \sim N(\mu, \sigma^2)$, probabilities involving X are computed by "standardizing." The standardized variable is $(X - \mu)/\sigma$. If X has a normal distribution with mean μ and standard deviation σ, then

$$Z = \frac{X - \mu}{\sigma},$$

has a standard normal distribution. Thus

$$\begin{aligned}\Pr(a \leq X \leq b) &= \Pr\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).\end{aligned}$$

Therefore,

$$\Pr(X \leq a) = \Phi\left(\frac{a-\mu}{\sigma}\right) \quad \text{and} \quad \Pr(X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right).$$

EXAMPLE 3.30

The time that it takes a driver to react to the brake lights on a decelerating vehicle is critical in avoiding rear-end collisions. It is suggested that reaction time for an in-traffic response to a brake signal from standard brake lights can be modelled with a normal distribution having mean value 1.25 s and standard deviation of 0.46 s. What is the probability that reaction time is between 1.00 and 1.75 s?

Solution. Let X denote reaction time, then standardizing gives

$$1 \leq X \leq 1.75$$

if and only if

$$\frac{1-1.25}{0.46} \leq \frac{x-1.25}{0.46} \leq \frac{1.75-1.25}{0.46}.$$

Thus

$$\begin{aligned}\Pr(1 \leq X \leq 1.75) &= \Pr\left(\frac{1-1.25}{0.46} \leq Z \leq \frac{1.75-1.25}{0.46}\right) \\ &= \Pr(-0.54 \leq Z \leq 1.09) = \Phi(1.09) - \Phi(-0.54) \\ &= 0.8621 - 0.2946 = 0.5675.\end{aligned}$$

◀

EXAMPLE 3.31

If the r.v $X \sim N(\mu, \sigma^2)$, such that $\Pr(X \leq 60) = 0.1$ and $\Pr(X > 90) = 0.05$. Find μ and σ^2.

Solution. Since $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$. Hence

$$0.1 = \Pr(X \leq 60) = \Pr\left(\frac{X-\mu}{\sigma} \leq \frac{60-\mu}{\sigma}\right) = \Pr\left(Z \leq \frac{60-\mu}{\sigma}\right).$$

From $\Phi(z)$ tables, we can find that $\frac{60-\mu}{\sigma} = -1.282$.

Similarly,

$$0.05 = \Pr(X > 90) = \Pr\left(\frac{X-\mu}{\sigma} > \frac{90-\mu}{\sigma}\right) = \Pr\left(Z > \frac{90-\mu}{\sigma}\right) = 1 - \Pr\left(Z \leq \frac{90-\mu}{\sigma}\right).$$

From $\Phi(z)$ tables, we can find that $\frac{90-\mu}{\sigma} = 1.645$. Then, it is easy to find that $\mu = 73.1$ and $\sigma = 10.2$. ◀

***Note*:**

1. If a r.v $X \sim N(0,1)$, then the r.v $Y = X^2 \sim \chi^2(1)$.
2. If a r.v $Z \sim N(0,1)$, then $\Pr(Z \leq -a) = \Pr(Z > a)$.
3. If a r.v $Z \sim N(0,1)$, then $\Pr(-a \leq Z \leq a) = 2\Pr(Z \leq a) - 1$

***Moment Generating Function*:** The mgf of a r.v $X \sim N(\mu, \sigma^2)$ is

$$M(t) = E[e^{tX}] = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

To prove that,

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{tx} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(x^2 - 2\mu x + \mu^2 - 2\sigma^2 tx)} dx,$$

take

$$\begin{aligned}x^2 - 2\mu x + \mu^2 - 2\sigma^2 tx &= x^2 - 2(\mu + \sigma^2 t)x + (\mu + \sigma^2 t)^2 - (\mu + \sigma^2 t)^2 + \mu^2 \\ &= [x - (\mu + \sigma^2 t)]^2 - \mu - 2\mu\sigma^2 t - \sigma^4 t^2 + \mu^2 \\ &= [x - (\mu + \sigma^2 t)]^2 - 2(\mu t + \frac{1}{2}\sigma^2 t^2)\sigma^2.\end{aligned}$$

Therefore,

$$M(t) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[x-(\mu+\sigma^2 t)]^2}\, e^{\mu t+\frac{1}{2}\sigma^2 t^2} dx = e^{\mu t+\frac{1}{2}\sigma^2 t^2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}[x-(\mu+\sigma^2 t)]^2}\, dx.$$

That gives $M(t) = e^{\mu t+\frac{1}{2}\sigma^2 t^2}$ since the integral is equal to one.

EXAMPLE 3.32

Let the r.v X has an mgf of $M(t) = e^{5t+2t^2}$. What is the distribution of X? Find $\Pr(3 \leq X \leq 7)$

Solution. We can rewrite the mgf as $M(t) = e^{5t+\frac{1}{2}4t^2}$, which means that $X \sim N(5, 4)$ with $\mu = 5$ and $\sigma^2 = 4$. To find the probability,

$$\begin{aligned}\Pr(3 \leq X \leq 7) &= \Pr\left(\frac{3-5}{2} \leq Z \leq \frac{7-5}{2}\right) = \Pr(-1 \leq Z \leq 1) = 2\Pr(Z \leq 1) - 1 \\ &= 2(0.841) - 1 = 1.682 - 1 = 0.682.\end{aligned}$$

◀