A Probability Course for the Actuaries
A Preparation for Exam P/1

Marcel B. Finan
Arkansas Tech University

November 2013 Syllabus
Last updated

July 8, 2014

In memory of my parents

August 1, 2008
January 7, 2009

Preface

The present manuscript is designed mainly to help students prepare for the Probability Exam (Exam P/1), the first actuarial examination administered by the Society of Actuaries. This examination tests a student's knowledge of the fundamental probability tools for quantitatively assessing risk. A thorough command of calculus is assumed.
More information about the exam can be found on the webpage of the Society of Actuaries www.soa.org.
Problems taken from previous exams provided by the Society of Actuaries will be indicated by the symbol ‡.
The flow of topics in the book follows very closely that of Ross's *A First Course in Probability*, 8th edition. Selected topics are chosen based on July 2013 exam syllabus as posted on the SOA website.
This manuscript can be used for personal use or class use, but not for commercial purposes. If you find any errors, I would appreciate hearing from you: mfinan@atu.edu
This manuscript is also suitable for a one semester course in an undergraduate course in probability theory. Answer keys to text problems are found at the end of the book.

Marcel B. Finan
Russellville, AR
August, 2013

Contents

A Review of Set Theory

The axiomatic approach to probability is developed using the foundation of set theory, and a quick review of the theory is in order. If you are familiar with set builder notation, Venn diagrams, and the basic operations on sets, (unions, intersections, and complements), then you have a good start on what we will need right away from set theory.
Set is the most basic term in mathematics. Some synonyms of a set are *class* or *collection*. In this chapter we introduce the concept of a set and its various operations and then study the properties of these operations.

Throughout this book, we assume that the reader is familiar with the following number systems:

• The set of all positive integers

$$\mathbb{N} = \{1, 2, 3, \cdots\}.$$

• The set of all integers

$$\mathbb{Z} = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}.$$

• The set of all rational numbers

$$\mathbb{Q} = \{\frac{a}{b} : a, b \in \mathbb{Z} \ with \ b \neq 0\}.$$

• The set $\mathbb{R}$ of all real numbers.

1 Basic Definitions

We define a **set** A as a collection of *well-defined* objects (called **elements** or **members** of A) such that for any given object x one can assert without dispute that either $x \in A$ (i.e., x belongs to A) or $x \notin A$ but not both.

Example 1.1
Which of the following is a well-defined set.
(a) The collection of good movies.
(b) The collection of right-handed individuals in Russellville.

Solution.
(a) The collection of good movies is not a well-defined set since the answer to the question: "Is *Les Miserables* a good movie?" may be subject to dispute.
(b) This collection is a well-defined set since a person is either left-handed or right-handed. Of course, we are ignoring those few who can use both hands ■

There are two different ways for representing a set. The first one is to list, without repetition, the elements of the set. For example, if A is the solution set to the equation $x^2 - 4 = 0$ then $A = \{-2, 2\}$. The other way to represent a set is to describe a property that characterizes the elements of the set. This is known as the **set-builder** representation of a set. For example, the set A above can be written as $A = \{x|x \text{ is an integer satisfying } x^2 - 4 = 0\}$.
We define the **empty** set, denoted by $\emptyset$, to be the set with no elements. A set which is not empty is called a **non-empty** set.

Example 1.2
List the elements of the following sets.
(a) $\{x|x \text{ is a real number such that } x^2 = 1\}$.
(b) $\{x|x \text{ is an integer such that } x^2 - 3 = 0\}$.

Solution.
(a) $\{-1, 1\}$.
(b) Since the only solutions to the given equation are $-\sqrt{3}$ and $\sqrt{3}$ and both are not integers, the set in question is the empty set ■

Example 1.3
Use a property to give a description of each of the following sets.
(a) $\{\text{a}, \text{e}, \text{i}, \text{o}, \text{u}\}$.
(b) $\{1, 3, 5, 7, 9\}$.

Solution.
(a) $\{x|x \text{ is a vowel}\}$.
(b) $\{n \in \mathbb{N}|n \text{ is odd and less than } 10 \}$ ■

The first arithmetic operation involving sets that we consider is the equality of two sets. Two sets A and B are said to be **equal** if and only if they contain the same elements. We write $A = B$. For non-equal sets we write $A \neq B$. In this case, the two sets do not contain the same elements.

Example 1.4
Determine whether each of the following pairs of sets are equal.
(a) $\{1, 3, 5\}$ and $\{5, 3, 1\}$.
(b) $\{\{1\}\}$ and $\{1, \{1\}\}$.

Solution.
(a) Since the order of listing elements in a set is irrelevant, $\{1, 3, 5\} = \{5, 3, 1\}$.
(b) Since one of the sets has exactly one member and the other has two, $\{\{1\}\} \neq \{1, \{1\}\}$ ■

In set theory, the number of elements in a set has a special name. It is called the **cardinality** of the set. We write $n(A)$ to denote the cardinality of the set A. If A has a finite cardinality we say that A is a **finite** set. Otherwise, it is called **infinite**. For example, $\mathbb{N}$ is an infinite set.
Can two infinite sets have the same cardinality? The answer is yes. If A and B are two sets (finite or infinite) and there is a **bijection** from A to B(i.e., a **one-to-one**[1] and **onto**[2] function) then the two sets are said to have the **same cardinality** and we write $n(A) = n(B)$.
If $n(A)$ is either finite or has the same cardinality as $\mathbb{N}$ then we say that A is **countable**. A set that is not countable is said to be **uncountable**.

Example 1.5
What is the cardinality of each of the following sets?
(a) $\emptyset$.

[1] A function $f : A \longmapsto B$ is a **one-to-one** function if $f(m) = f(n)$ implies $m = n$, where $m, n \in A$.
[2] A function $f : A \longmapsto B$ is an **onto** function if for every $b \in B$, there is an $a \in A$ such that $b = f(a)$.

(b) $\{\emptyset\}$.
(c) $A = \{a, \{a\}, \{a, \{a\}\}\}$.

Solution.
(a) $n(\emptyset) = 0$.
(b) This is a set consisting of one element $\emptyset$. Thus, $n(\{\emptyset\}) = 1$.
(c) $n(A) = 3$ ∎

Example 1.6
(a) Show that the set $A = \{a_1, a_2. \cdots, a_n, \cdots\}$ is countable.
(b) Let A be the set of all infinite sequences of the digits 0 and 1. Show that A is uncountable.

Solution.
(a) One can easily verify that the map $f : \mathbb{N} \longmapsto A$ defined by $f(n) = a_n$ is a bijection.
(b) We will argue by contradiction. So suppose that A is countable with elements $a_1, a_2, \cdots$. where each a_i is an infinite sequence of the digits 0 and 1. Let a be the infinite sequence with the first digit of 0 or 1 different from the first digit of a_1, the second digit of 0 or 1 different from the second digit of $a_2, \cdots$, the n^{th} digit is different from the n^{th} digit of a_n, etc. Thus, a is an infinite sequence of the digits 0 and 1 which is not in A, a contradiction. Hence, A is uncountable ∎

Now, one compares numbers using inequalities. The corresponding notion for sets is the concept of a subset: Let A and B be two sets. We say that A is a **subset** of B, denoted by $A \subseteq B$, if and only if every element of A is also an element of B. If there exists an element of A which is not in B then we write $A \not\subseteq B$.
For any set A we have $\emptyset \subseteq A \subseteq A$. That is, every set has at least two subsets. Also, keep in mind that the empty set is a subset of any set.

Example 1.7
Suppose that $A = \{2, 4, 6\}$, $B = \{2, 6\}$, and $C = \{4, 6\}$. Determine which of these sets are subsets of which other of these sets.

Solution.
$B \subseteq A$ and $C \subseteq A$ ∎

If sets A and B are represented as regions in the plane, relationships between A and B can be represented by pictures, called **Venn diagrams.**

Example 1.8
Represent $A \subseteq B \subseteq C$ using Venn diagram.

Solution.
The Venn diagram is given in Figure 1.1 ■

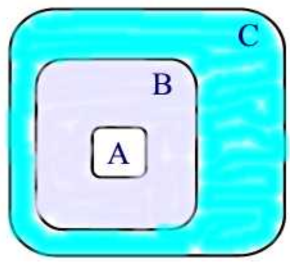


Figure 1.1

Let A and B be two sets. We say that A is a **proper** subset of B, denoted by $A \subset B$, if $A \subseteq B$ and $A \neq B$. Thus, to show that A is a proper subset of B we must show that every element of A is an element of B and there is an element of B which is not in A.

Example 1.9
Order the sets of numbers: $\mathbb{Z}, \mathbb{R}, \mathbb{Q}, \mathbb{N}$ using $\subset$

Solution.
$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ ■

Example 1.10
Determine whether each of the following statements is true or false.
(a) $x \in \{x\}$ (b) $\{x\} \subseteq \{x\}$ (c) $\{x\} \in \{x\}$
(d) $\{x\} \in \{\{x\}\}$ (e) $\emptyset \subseteq \{x\}$ (f) $\emptyset \in \{x\}$

Solution.
(a) True (b) True (c) False since $\{x\}$ is a set consisting of a single element x and so $\{x\}$ is not a member of this set (d) True (e) True (f) False since $\{x\}$ does not have $\emptyset$ as a listed member ■

Now, the collection of all subsets of a set A is of importance. We denote this set by $\mathcal{P}$(A) and we call it the **power set** of A.

Example 1.11
Find the power set of $A = \{a, b, c\}$.

Solution.

$$\mathcal{P}(\mathrm{A}) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\} \blacksquare$$

We conclude this section, by introducing the concept of **mathematical induction**: We want to prove that some statement $P(n)$ is true for any nonnegative integer $n \geq n_0$. The steps of mathematical induction are as follows:

(i) (Basis of induction) Show that $P(n_0)$ is true.
(ii) (Induction hypothesis) Assume $P(n_0), P(n_0 + 1), \cdots, P(n)$ are true.
(iii) (Induction step) Show that $P(n + 1)$ is true.

Example 1.12
(a) Use induction to show that if $n(A) = n$ then $n(\mathcal{P}(\mathrm{A})) = 2^n$, where $n \geq 0$ and $n \in \mathbb{N}$.
(b) If $\mathcal{P}(\mathrm{A})$ has 256 elements, how many elements are there in A?

Solution.
(a) We apply induction to prove the claim. If $n = 0$ then $A = \emptyset$ and in this case $\mathcal{P}(\mathrm{A}) = \{\emptyset\}$. Thus, $n(\mathcal{P}(\mathrm{A})) = 1 = 2^0$. As induction hypothesis, suppose that if $n(A) = n$ then $n(\mathcal{P}(\mathrm{A})) = 2^n$. Let $B = \{a_1, a_2, \cdots, a_n, a_{n+1}\}$. Then $\mathcal{P}(\mathrm{B})$ consists of all subsets of $\{a_1, a_2, \cdots, a_n\}$ together with all subsets of $\{a_1, a_2, \cdots, a_n\}$ with the element a_{n+1} added to them. Hence, $n(\mathcal{P}(\mathrm{B})) = 2^n + 2^n = 2 \cdot 2^n = 2^{n+1}$.
(b) Since $n(\mathcal{P}(\mathrm{A})) = 256 = 2^8$, by (a) we have $n(A) = 8$ $\blacksquare$

Example 1.13
Use induction to show that $\displaystyle\sum_{i=1}^{n}(2i - 1) = n^2,\ n \in \mathbb{N}$.

Solution.
If $n = 1$ we have $1^2 = 2(1) - 1 = \displaystyle\sum_{i=1}^{1}(2i - 1)$. Suppose that the result is true for up to n. We will show that it is true for $n + 1$. Indeed, $\displaystyle\sum_{i=1}^{n+1}(2i - 1) = \sum_{i=1}^{n}(2i - 1) + 2(n + 1) - 1 = n^2 + 2n + 2 - 1 = (n + 1)^2$ $\blacksquare$

Practice Problems

Problem 1.1
Consider the experiment of rolling a die. List the elements of the set $A = \{x : x \text{ shows a face with prime number}\}$. Recall that a **prime** number is a number with only two different divisors: 1 and the number itself.

Problem 1.2
Consider the random experiment of tossing a coin three times.
(a) Let S be the collection of all outcomes of this experiment. List the elements of S. Use H for head and T for tail.
(b) Let E be the subset of S with more than one tail. List the elements of E.
(c) Suppose $F = \{THH, HTH, HHT, HHH\}$. Write F in set-builder notation.

Problem 1.3
Consider the experiment of tossing a coin three times. Let E be the collection of outcomes with at least one head and F the collection of outcomes of more than one head. Compare the two sets E and F.

Problem 1.4
A hand of 5 cards is dealt from a deck of 52 cards. Let E be the event that the hand contains 5 aces. List the elements of E.

Problem 1.5
Prove the following properties:
(a) Reflexive Property: $A \subseteq A$.
(b) Antisymmetric Property: If $A \subseteq B$ and $B \subseteq A$ then $A = B$.
(c) Transitive Property: If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.

Problem 1.6
Prove by using mathematical induction that

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}, \quad n \in \mathbb{N}.$$

Problem 1.7
Prove by using mathematical induction that

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}, \quad n \in \mathbb{N}.$$

Problem 1.8
Use induction to show that $(1+x)^n \geq 1+nx$ for all $n \in \mathbb{N}$, where $x > -1$.

Problem 1.9
Use induction to show that

$$1 + a + a^2 + \cdots + a^{n-1} = \frac{1-a^n}{1-a}.$$

Problem 1.10
Subway prepared 60 4-inch sandwiches for a birthday party. Among these sandwiches, 45 of them had tomatoes, 30 had both tomatoes and onions, and 5 had neither tomatoes nor onions. Using a Venn diagram, how many sandwiches did he make with
(a) tomatoes or onions?
(b) onions?
(c) onions but not tomatoes?

Problem 1.11
A camp of international students has 110 students. Among these students,

75	speak english,
52	speak spanish,
50	speak french,
33	speak english and spanish,
30	speak english and french,
22	speak spanish and french,
13	speak all three languages.

How many students speak
(a) english and spanish, but not french,
(b) neither english, spanish, nor french,
(c) french, but neither english nor spanish,
(d) english, but not spanish,
(e) only one of the three languages,
(f) exactly two of the three languages.

Problem 1.12
An experiment consists of the following two stages:
(1) a fair coin is tossed

(2) if the coin shows a head, then a fair die is rolled; otherwise, the coin is flipped again.
An outcome of this experiment is a pair of the form (outcome from stage 1, outcome from stage 2). Let S be the collection of all outcomes. List the elements of S and then find the cardinality of S.

Problem 1.13
Show that the function $f : \mathbb{R} \longmapsto \mathbb{R}$ defined by $f(x) = 3x + 5$ is one-to-one and onto.

Problem 1.14
Find $n(A)$ if $n(\mathcal{P}(A)) = 32$.

Problem 1.15
Consider the function $f : \mathbb{N} \longmapsto \mathbb{Z}$ defined by

$$f(n) = \begin{cases} \frac{n}{2}, & \text{if } n \text{ is even} \\ -\frac{n-1}{2}, & \text{if } n \text{ is odd.} \end{cases}$$

(a) Show that $f(n) = f(m)$ cannot happen if n and m have different parity, i.e., either both are even or both are odd..
(b) Show that $\mathbb{Z}$ is countable.

Problem 1.16
Let A be a non-empty set and $f : A \longmapsto \mathcal{P}(A)$ be any function. Let $B = \{a \in A | a \not\in f(a)\}$. Clearly, $B \in \mathcal{P}(A)$. Show that there is no $b \in A$ such that $f(b) = B$. Hence, there is no onto map from A to $\mathcal{P}(A)$.

Problem 1.17
Use the previous problem to show that $\mathcal{P}(\mathbb{N})$ is uncountable.

2 Set Operations

In this section we introduce various operations on sets and study the properties of these operations.

Complements
If U is a given set whose subsets are under consideration, then we call U a **universal set.** Let U be a universal set and A, B be two subsets of U. The **absolute complement** of A (See Figure 2.1(I)) is the set

$$A^c = \{x \in U | x \notin A\}.$$

Example 2.1
Find the complement of $A = \{1, 2, 3\}$ if $U = \{1, 2, 3, 4, 5, 6\}$.

Solution.
From the definition, $A^c = \{4, 5, 6\}$ ■

The **relative complement** of A with respect to B (See Figure 2.1(II)) is the set

$$B - A = \{x \in U | x \in B \text{ and } x \notin A\}.$$

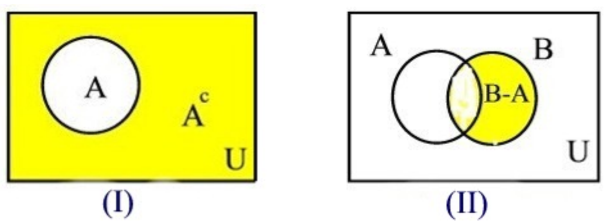


Figure 2.1

Example 2.2
Let $A = \{1, 2, 3\}$ and $B = \{\{1, 2\}, 3\}$. Find $A - B$.

Solution.
The elements of A that are not in B are 1 and 2. That is, $A - B = \{1, 2\}$ ■

Union and Intersection
Given two sets A and B. The **union** of A and B is the set

$$A \cup B = \{x | x \in A \text{ or } x \in B\}$$

where the 'or' is inclusive.(See Figure 2.2(a))

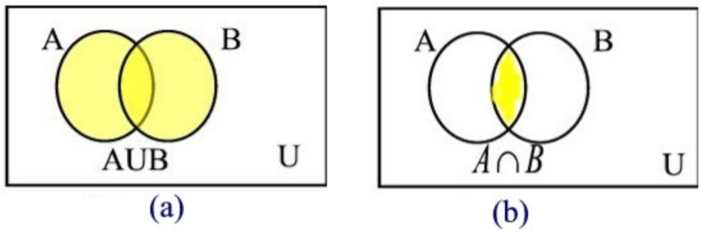


Figure 2.2

The above definition can be extended to more than two sets. More precisely, if $A_1, A_2, \cdots$, are sets then

$$\bigcup_{n=1}^{\infty} A_n = \{x | x \in A_i \ for \ some \ i \in \mathbb{N}\}.$$

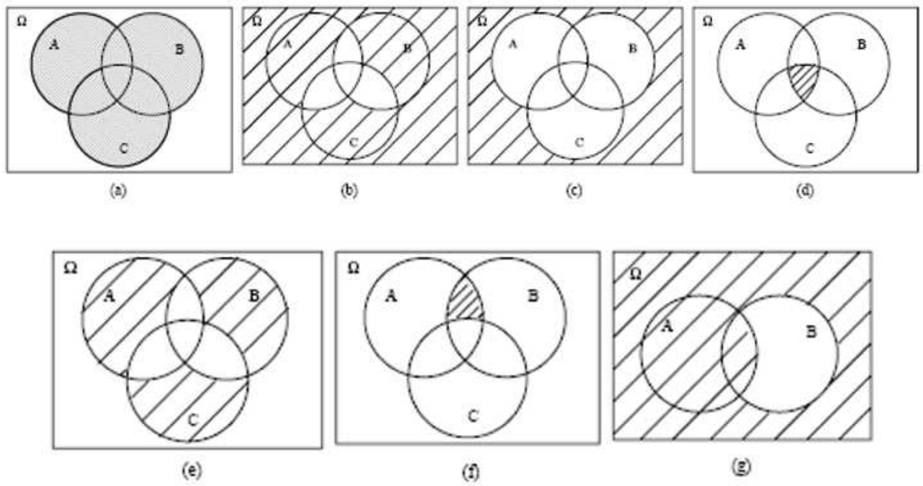
The **intersection** of A and B is the set (See Figure 2.2(b))

$$A \cap B = \{x | x \in A \ and \ x \in B\}.$$

Example 2.3
Express each of the following events in terms of the events A, B, and C as well as the operations of complementation, union and intersection:
(a) at least one of the events A, B, C occurs;
(b) at most one of the events A, B, C occurs;
(c) none of the events A, B, C occurs;
(d) all three events A, B, C occur;
(e) exactly one of the events A, B, C occurs;
(f) events A and B occur, but not C;
(g) either event A occurs or, if not, then B also does not occur.
In each case draw the corresponding Venn diagram.

Solution.
(a) $A \cup B \cup C$
(b) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C) \cup (A^c \cap B^c \cap C^c)$
(c) $(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$
(d) $A \cap B \cap C$
(e) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$

(f) $A \cap B \cap C^c$
(g) $A \cup (A^c \cap B^c)$ ■



Example 2.4
Translate the following set-theoretic notation into event language. For example, "$A \cup B$" means "A or B occurs".
(a) $A \cap B$
(b) $A - B$
(c) $A \cup B - A \cap B$
(d) $A - (B \cup C)$
(e) $A \subset B$
(f) $A \cap B = \emptyset$

Solution.
(a) A and B occur
(b) A occurs and B does not occur
(c) A or B, but not both, occur
(d) A occurs, and B and C do not occur
(e) if A occurs, then B occurs but if B occurs then A need not occur.
(f) if A occurs, then B does not occur or if B occurs then A does not occur ■

Example 2.5
Find a simpler expression of $[(A \cup B) \cap (A \cup C) \cap (B^c \cap C^c)]$ assuming all three sets A, B, C intersect.

Solution.
Using a Venn diagram one can easily see that $[(A \cup B) \cap (A \cup C) \cap (B^c \cap C^c)] = A - [A \cap (B \cup C)] = A - B \cup C$ ■

If $A \cap B = \emptyset$ we say that A and B are **disjoint** sets.

Example 2.6
Let A and B be two non-empty sets. Write A as the union of two disjoint sets.

Solution.
Using a Venn diagram one can easily see that $A \cap B$ and $A \cap B^c$ are disjoint sets such that $A = (A \cap B) \cup (A \cap B^c)$ ■

Example 2.7
In a junior league tennis tournament, teams play 20 games. Let A denote the event that Team Blazers wins 15 or more games in the tournament. Let B be the event that the Blazers win less than 10 games and C be the event that they win between 8 to 16 games. The Blazers can win at most 20 games. Using words, what do the following events represent?
(a) $A \cup B$ and $A \cap B$.
(b) $A \cup C$ and $A \cap C$.
(c) $B \cup C$ and $B \cap C$.
(d) A^c, B^c, and C^c.

Solution.
(a) $A \cup B$ is the event that the Blazers win 15 or more games or win 9 or less games. $A \cap B$ is the empty set, since the Blazers cannot win 15 or more games and have less than 10 wins at the same time. Therefore, event A and event B are disjoint.
(b) $A \cup C$ is the event that the Blazers win at least 8 games. $A \cap C$ is the event that the Blazers win 15 or 16 games.
(c) $B \cup C$ is the event that the Blazers win at most 16 games. $B \cap C$ is the event that the Blazers win 8 or 9 games.
(d) A^c is the event that the Blazers win 14 or fewer games. B^c is the event that the Blazers win 10 or more games. C^c is the event that the Blazers win fewer than 8 or more than 16 games ■

Given the sets $A_1, A_2, \cdots$, we define

$$\bigcap_{n=1}^{\infty} A_n = \{x | x \in A_i \ for\ all\ i \in \mathbb{N}\}.$$

Example 2.8

For each positive integer n we define $A_n = \{n\}$. Find $\bigcap_{n=1}^{\infty} A_n$.

Solution.

Clearly, $\bigcap_{n=1}^{\infty} A_n = \emptyset$ ■

Remark 2.1

Note that the Venn diagrams of $A \cap B$ and $A \cup B$ show that $A \cap B = B \cap A$ and $A \cup B = B \cup A$. That is, $\cup$ and $\cap$ are commutative laws.

The following theorem establishes the distributive laws of sets.

Theorem 2.1

If A, B, and C are subsets of U then
(a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
(b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Proof.

See Problem 2.15 ■

Remark 2.2

Note that since $\cap$ and $\cup$ are commutative operations, we have $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ and $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.

The following theorem presents the relationships between $(A \cup B)^c, (A \cap B)^c, A^c$ and B^c.

Theorem 2.2 *(De Morgan's Laws)*

Let A and B be subsets of U. We have
(a) $(A \cup B)^c = A^c \cap B^c$.
(b) $(A \cap B)^c = A^c \cup B^c$.

Proof.
We prove part (a) leaving part(b) as an exercise for the reader.
(a) Let $x \in (A \cup B)^c$. Then $x \in U$ and $x \notin A \cup B$. Hence, $x \in U$ and ($x \notin A$ and $x \notin B$). This implies that ($x \in U$ and $x \notin A$) and ($x \in U$ and $x \notin B$). It follows that $x \in A^c \cap B^c$.
Conversely, let $x \in A^c \cap B^c$. Then $x \in A^c$ and $x \in B^c$. Hence, $x \notin A$ and $x \notin B$ which implies that $x \notin (A \cup B)$. Hence, $x \in (A \cup B)^c$ ■

Remark 2.3
De Morgan's laws are valid for any countable number of sets. That is

$$\left(\bigcup_{n=1}^{\infty} A_n\right)^c = \bigcap_{n=1}^{\infty} A_n^c$$

and

$$\left(\bigcap_{n=1}^{\infty} A_n\right)^c = \bigcup_{n=1}^{\infty} A_n^c$$

Example 2.9
An assisted living agency advertises its program through videos and booklets. Let U be the set of people solicited for the agency program. All participants were given a chance to watch a video and to read a booklet describing the program. Let V be the set of people who watched the video, B the set of people who read the booklet, and C the set of people who decided to enroll in the program.
(a) Describe with set notation: "The set of people who did not see the video or read the booklet but who still enrolled in the program"
(b) Rewrite your answer using De Morgan's law and and then restate the above.

Solution.
(a) $(V \cup B)^c \cap C$.
(b) $(V \cup B)^c \cap C = V^c \cap B^c \cap C =$ the set of people who did not watch the video, did not read the booklet, but did enroll ■

If $A_i \cap A_j = \emptyset$ for all $i \neq j$ then we say that the sets in the collection $\{A_n\}_{n=1}^{\infty}$ are **pairwise disjoint.**

Example 2.10
Find three sets A, B, and C that are not pairwise disjoint but $A \cap B \cap C = \emptyset$.

Solution.
One example is $A = B = \{1\}$ and $C = \emptyset$ ■

Example 2.11
Find sets $A_1, A_2, \cdots$ that are pairwise disjoint and $\bigcap_{n=1}^{\infty} A_n = \emptyset$.

Solution.
For each positive integer n, let $A_n = \{n\}$ ■

Example 2.12
Throw a pair of fair dice. Let A be the event the total is 5, B the event the total is even, and C the event the total is divisible by 9. Show that A, B, and C are pairwise disjoint.

Solution.
We have

$$
\begin{aligned}
A =& \{(1,4),(2,3),(3,2),(4,1)\} \\
B =& \{(1,1),(1,3),(1,5),(2,2),(2,4),(2,6)(3,1),(3,3),(3,5),(4,2), \\
& (4,4),(4,6),(5,1),(5,3),(5,5),(6,2),(6,4),(6,6)\} \\
C =& \{(3,6),(4,5),(5,4),(6,3)\}.
\end{aligned}
$$

Clearly, $A \cap B = A \cap C = B \cap C = \emptyset$ ■

Next, we establish the following rule of counting.

Theorem 2.3 *(Inclusion-Exclusion Principle)*
Suppose A and B are finite sets. Then
(a) $n(A \cup B) = n(A) + n(B) - n(A \cap B)$.
(b) If $A \cap B = \emptyset$, then $n(A \cup B) = n(A) + n(B)$.
(c) If $A \subseteq B$, then $n(A) \leq n(B)$.

Proof.
(a) Indeed, $n(A)$ gives the number of elements in A including those that are common to A and B. The same holds for $n(B)$. Hence, $n(A) + n(B)$ includes

twice the number of common elements. Therefore, to get an accurate count of the elements of $A \cup B$, it is necessary to subtract $n(A \cap B)$ from $n(A)+n(B)$. This establishes the result.
(b) If A and B are disjoint then $n(A \cap B) = 0$ and by (a) we have $n(A \cup B) = n(A) + n(B)$.
(c) If A is a subset of B then the number of elements of A cannot exceed the number of elements of B. That is, $n(A) \leq n(B)$ ■

Example 2.13
The State Department interviewed 35 candidates for a diplomatic post in Algeria; 25 speak arabic, 28 speak french, and 2 speak neither languages. How many speak both languages?

Solution.
Let F be the group of applicants that speak french, A those who speak arabic. Then $F \cap A$ consists if those who speak both languages. By the Inclusion-Exclusion Principle we have $n(F \cup A) = n(F) + n(A) - n(F \cap A)$. That is, $33 = 28 + 25 - n(F \cap A)$. Solving for $n(F \cap A)$ we find $n(F \cap A) = 20$ ■

Cartesian Product
The notation (a, b) is known as an **ordered pair** of elements and is defined by $(a, b) = \{\{a\}, \{a, b\}\}$.
The **Cartesian product** of two sets A and B is the set

$$A \times B = \{(a, b) | a \in A, b \in B\}.$$

The idea can be extended to products of any number of sets. Given n sets $A_1, A_2, \cdots, A_n$ the Cartesian product of these sets is the set

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \cdots, a_n) : a_1 \in A_1, a_2 \in A_2, \cdots, a_n \in A_n\}$$

Example 2.14
Consider the experiment of tossing a fair coin n times. Represent the sample space as a Cartesian product.

Solution.
If S is the sample space then $S = S_1 \times S_2 \times \cdots \times S_n$, where S_i, $1 \leq i \leq n$, is the set consisting of the two outcomes H=head and T = tail ■

The following theorem is a tool for finding the cardinality of the Cartesian product of two finite sets.

Theorem 2.4
Given two finite sets A and B. Then

$$n(A \times B) = n(A) \cdot n(B).$$

Proof.
Suppose that $A = \{a_1, a_2, \cdots, a_n\}$ and $B = \{b_1, b_2, \cdots, b_m\}$. Then

$$\begin{aligned} A \times B = \{&(a_1, b_1), (a_1, b_2), \cdots, (a_1, b_m),\\ &(a_2, b_1), (a_2, b_2), \cdots, (a_2, b_m),\\ &(a_3, b_1), (a_3, b_2), \cdots, (a_3, b_m),\\ &\vdots\\ &(a_n, b_1), (a_n, b_2), \cdots, (a_n, b_m)\} \end{aligned}$$

Thus, $n(A \times B) = n \cdot m = n(A) \cdot n(B)$ ■

Remark 2.4
By induction, the previous result can be extended to any finite number of sets.

Example 2.15
What is the total number of outcomes of tossing a fair coin n times.

Solution.
If S is the sample space then $S = S_1 \times S_2 \times \cdots \times S_n$ where $S_i,\ 1 \le i \le n$, is the set consisting of the two outcomes H=head and T = tail. By the previous theorem, $n(S) = 2^n$ ■

Practice Problems

Problem 2.1
Let A and B be any two sets. Use Venn diagrams to show that $B = (A \cap B) \cup (A^c \cap B)$ and $A \cup B = A \cup (A^c \cap B)$.

Problem 2.2
Show that if $A \subseteq B$ then $B = A \cup (A^c \cap B)$. Thus, B can be written as the union of two disjoint sets.

Problem 2.3 ‡
A survey of a group's viewing habits over the last year revealed the following information

(i) 28% watched gymnastics
(ii) 29% watched baseball
(iii) 19% watched soccer
(iv) 14% watched gymnastics and baseball
(v) 12% watched baseball and soccer
(vi) 10% watched gymnastics and soccer
(vii) 8% watched all three sports.

Represent the statement "the group that watched none of the three sports during the last year" using operations on sets.

Problem 2.4
An urn contains 10 balls: 4 red and 6 blue. A second urn contains 16 red balls and an unknown number of blue balls. A single ball is drawn from each urn. For $i = 1, 2$, let R_i denote the event that a red ball is drawn from urn i and B_i the event that a blue ball is drawn from urn i. Show that the sets $R_1 \cap R_2$ and $B_1 \cap B_2$ are disjoint.

Problem 2.5 ‡
An auto insurance has 10,000 policyholders. Each policyholder is classified as

(i) young or old;
(ii) male or female;
(iii) married or single.

Of these policyholders, 3,000 are young, 4,600 are male, and 7,000 are married. The policyholders can also be classified as 1,320 young males, 3,010 married males, and 1,400 young married persons. Finally, 600 of the policyholders are young married males.
How many of the company's policyholders are young, female, and single?

Problem 2.6 ‡
A marketing survey indicates that 60% of the population owns an automobile, 30% owns a house, and 20% owns both an automobile and a house. What percentage of the population owns an automobile or a house, but not both?

Problem 2.7 ‡
35% of visits to a primary care physicians (PCP) office results in neither lab work nor referral to a specialist. Of those coming to a PCPs office, 30% are referred to specialists and 40% require lab work.
What percentage of visit to a PCPs office results in both lab work and referral to a specialist?

Problem 2.8
In a universe U of 100, let A and B be subsets of U such that $n(A \cup B) = 70$ and $n(A \cup B^c) = 90$. Determine $n(A)$.

Problem 2.9 ‡
An insurance company estimates that 40% of policyholders who have only an auto policy will renew next year and 60% of policyholders who have only a homeowners policy will renew next year. The company estimates that 80% of policyholders who have both an auto and a homeowners policy will renew at least one of those policies next year. Company records show that 65% of policyholders have an auto policy, 50% of policyholders have a homeowners policy, and 15% of policyholders have both an auto and a homeowners policy. Using the company's estimates, calculate the percentage of policyholders that will renew at least one policy next year.

Problem 2.10
Show that if A, B, and C are subsets of a universe U then

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

Problem 2.11
In a survey on popsicle flavor preferences of kids aged 3-5, it was found that
• 22 like strawberry.
• 25 like blueberry.
• 39 like grape.
• 9 like blueberry and strawberry.
• 17 like strawberry and grape.
• 20 like blueberry and grape.
• 6 like all flavors.
• 4 like none.

How many kids were surveyed?

Problem 2.12
Let A, B, and C be three subsets of a universe U with the following properties: $n(A) = 63, n(B) = 91, n(C) = 44, n(A \cap B) = 25, n(A \cap C) = 23, n(C \cap B) = 21, n(A \cup B \cup C) = 139$. Find $n(A \cap B \cap C)$.

Problem 2.13
Fifty students living in a college dormitory were registering for classes for the fall semester. The following were observed:
• 30 registered in a math class,
• 18 registered in a history class,
• 26 registered in a computer class,
• 9 registered in both math and history classes,
• 16 registered in both math and computer classes,
• 8 registered in both history and computer classes,
• 47 registered in at least one of the three classes.
(a) How many students did not register in any of these classes ?
(b) How many students registered in all three classes?

Problem 2.14 ‡
A doctor is studying the relationship between blood pressure and heartbeat abnormalities in her patients. She tests a random sample of her patients and notes their blood pressures (high, low, or normal) and their heartbeats (regular or irregular). She finds that:

(i) 14% have high blood pressure.
(ii) 22% have low blood pressure.
(iii) 15% have an irregular heartbeat.
(iv) Of those with an irregular heartbeat, one-third have high blood pressure.
(v) Of those with normal blood pressure, one-eighth have an irregular heartbeat.

What portion of the patients selected have a regular heartbeat and low blood pressure?

Problem 2.15
Prove: If A, B, and C are subsets of U then
(a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
(b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Problem 2.16
Translate the following verbal description of events into set theoretic notation. For example, "A or B occurs, but not both" corresponds to the set $A \cup B - A \cap B$.
(a) A occurs whenever B occurs.
(b) If A occurs, then B does not occur.
(c) Exactly one of the events A and B occurs.
(d) Neither A nor B occur.

Problem 2.17 ‡
A survey of 100 TV watchers revealed that over the last year:
i) 34 watched CBS.
ii) 15 watched NBC.
iii) 10 watched ABC.
iv) 7 watched CBS and NBC.
v) 6 watched CBS and ABC.
vi) 5 watched NBC and ABC.
vii) 4 watched CBS, NBC, and ABC.
viii) 18 watched HGTV and of these, none watched CBS, NBC, or ABC.
Calculate how many of the 100 TV watchers did not watch any of the four channels (CBS, NBC, ABC or HGTV).

Counting and Combinatorics

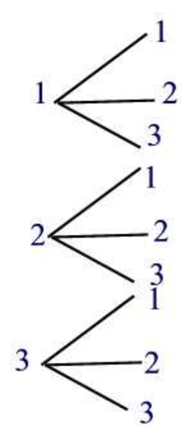
The major goal of this chapter is to establish several (combinatorial) techniques for counting large finite sets without actually listing their elements. These techniques provide effective methods for counting the size of events, an important concept in probability theory.

3 The Fundamental Principle of Counting

Sometimes one encounters the question of listing all the outcomes of a certain experiment. One way for doing that is by constructing a so-called **tree diagram.**

Example 3.1
List all two-digit numbers that can be constructed from the digits 1,2, and 3.

Solution.



The different numbers are $\{11, 12, 13, 21, 22, 23, 31, 32, 33\}$ ∎

Of course, trees are manageable as long as the number of outcomes is not large. If there are many stages to an experiment and several possibilities at each stage, the tree diagram associated with the experiment would become too large to be manageable. For such problems the counting of the outcomes is simplified by means of algebraic formulas. The commonly used formula is the **Fundamental Principle of Counting**, also known as the **multiplication rule of counting**, which states:

Theorem 3.1
If a choice consists of k steps, of which the first can be made in n_1 ways, for each of these the second can be made in n_2 ways,$\cdots$, and for each of these the k^{th} can be made in n_k ways, then the whole choice can be made in $n_1 \cdot n_2 \cdot \cdots n_k$ ways.

Proof.
In set-theoretic term, we let S_i denote the set of outcomes for the i^{th} task, $i = 1, 2, \cdots, k$. Note that $n(S_i) = n_i$. Then the set of outcomes for the entire job is the Cartesian product $S_1 \times S_2 \times \cdots \times S_k = \{(s_1, s_2, \cdots, s_k) : s_i \in S_i, 1 \le i \le k\}$. Thus, we just need to show that

$$n(S_1 \times S_2 \times \cdots \times S_k) = n(S_1) \cdot n(S_2) \cdots n(S_k).$$

The proof is by induction on $k \ge 2$.

Basis of Induction
This is just Theorem 2.4.
Induction Hypothesis
Suppose

$$n(S_1 \times S_2 \times \cdots \times S_k) = n(S_1) \cdot n(S_2) \cdots n(S_k).$$

Induction Conclusion
We must show

$$n(S_1 \times S_2 \times \cdots \times S_{k+1}) = n(S_1) \cdot n(S_2) \cdots n(S_{k+1}).$$

To see this, note that there is a one-to-one correspondence between the sets $S_1 \times S_2 \times \cdots \times S_{k+1}$ and $(S_1 \times S_2 \times \cdots S_k) \times S_{k+1}$ given by $f(s_1, s_2, \cdots, s_k, s_{k+1}) =$

$((s_1, s_2, \cdots, s_k), s_{k+1})$. Thus, $n(S_1 \times S_2 \times \cdots \times S_{k+1}) = n((S_1 \times S_2 \times \cdots S_k) \times S_{k+1}) = n(S_1 \times S_2 \times \cdots S_k)n(S_{k+1})$ (by Theorem 2.4). Now, applying the induction hypothesis gives

$$n(S_1 \times S_2 \times \cdots S_k \times S_{k+1}) = n(S_1) \cdot n(S_2) \cdots n(S_{k+1}) \blacksquare$$

Example 3.2
The following three factors were considered in the study of the effectivenenss of a certain cancer treatment:

(i) Medicine $(A_1, A_2, A_3, A_4, A_5)$
(ii) Dosage Level (Low, Medium, High)
(iii) Dosage Frequency (1,2,3,4 times/day)

Find the number of ways that a cancer patient can be given the medecine?

Solution.
The choice here consists of three stages, that is, $k = 3$. The first stage, can be made in $n_1 = 5$ different ways, the second in $n_2 = 3$ different ways, and the third in $n_3 = 4$ ways. Hence, the number of possible ways a cancer patient can be given medecine is $n_1 \cdot n_2 \cdot n_3 = 5 \cdot 3 \cdot 4 = 60$ different ways $\blacksquare$

Example 3.3
How many license-plates with 3 letters followed by 3 digits exist?

Solution.
A 6-step process: (1) Choose the first letter, (2) choose the second letter, (3) choose the third letter, (4) choose the first digit, (5) choose the second digit, and (6) choose the third digit. Every step can be done in a number of ways that does not depend on previous choices, and each license plate can be specified in this manner. So there are $26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 = 17,576,000$ ways $\blacksquare$

Example 3.4
How many numbers in the range 1000 - 9999 have no repeated digits?

Solution.
A 4-step process: (1) Choose first digit, (2) choose second digit, (3) choose third digit, (4) choose fourth digit. Every step can be done in a number of ways that does not depend on previous choices, and each number can be specified in this manner. So there are $9 \cdot 9 \cdot 8 \cdot 7 = 4,536$ ways $\blacksquare$

Example 3.5
How many license-plates with 3 letters followed by 3 digits exist if exactly one of the digits is 1?

Solution.
In this case, we must pick a place for the 1 digit, and then the remaining digit places must be populated from the digits $\{0, 2, \cdots 9\}$. A 6-step process: (1) Choose the first letter, (2) choose the second letter, (3) choose the third letter, (4) choose which of three positions the 1 goes, (5) choose the first of the other digits, and (6) choose the second of the other digits. Every step can be done in a number of ways that does not depend on previous choices, and each license plate can be specified in this manner. So there are $26 \cdot 26 \cdot 26 \cdot 3 \cdot 9 \cdot 9 = 4,270,968$ ways ■

Practice Problems

Problem 3.1
If each of the 10 digits 0-9 is chosen at random, how many ways can you choose the following numbers?
(a) A two-digit code number, repeated digits permitted.
(b) A three-digit identification card number, for which the first digit cannot be a 0. Repeated digits permitted.
(c) A four-digit bicycle lock number, where no digit can be used twice.
(d) A five-digit zip code number, with the first digit not zero. Repeated digits permitted.

Problem 3.2
(a) If eight cars are entered in a race and three finishing places are considered, how many finishing orders can they finish? Assume no ties.
(b) If the top three cars are Buick, Honda, and BMW, in how many possible orders can they finish?

Problem 3.3
You are taking 2 shirts(white and red) and 3 pairs of pants (black, blue, and gray) on a trip. How many different choices of outfits do you have?

Problem 3.4
A Poker club has 10 members. A president and a vice-president are to be selected. In how many ways can this be done if everyone is eligible?

Problem 3.5
In a medical study, patients are classified according to whether they have regular (RHB) or irregular heartbeat (IHB) and also according to whether their blood pressure is low (L), normal (N), or high (H). Use a tree diagram to represent the various outcomes that can occur.

Problem 3.6
If a travel agency offers special weekend trips to 12 different cities, by air, rail, bus, or sea, in how many different ways can such a trip be arranged?

Problem 3.7
If twenty different types of wine are entered in wine-tasting competition, in how many different ways can the judges award a first prize and a second prize?

Problem 3.8
In how many ways can the 24 members of a faculty senate of a college choose a president, a vice-president, a secretary, and a treasurer?

Problem 3.9
Find the number of ways in which four of ten new novels can be ranked first, second, third, and fourth according to their figure sales for the first three months.

Problem 3.10
How many ways are there to seat 8 people, consisting of 4 couples, in a row of seats (8 seats wide) if all couples are to get adjacent seats?

4 Permutations

Consider the following problem: In how many ways can 8 horses finish in a race (assuming there are no ties)? We can look at this problem as a decision consisting of 8 steps. The first step is the possibility of a horse to finish first in the race, the second step is the possibility of a horse to finish second, $\cdots$, the 8^{th} step is the possibility of a horse to finish 8^{th} in the race. Thus, by the Fundamental Principle of Counting there are

$$8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 40,320 \text{ ways.}$$

This problem exhibits an example of an ordered arrangement, that is, the order the objects are arranged is important. Such an ordered arrangement is called a **permutation**. Products such as $8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ can be written in a shorthand notation called factorial. That is, $8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 8!$ (read "8 factorial"). In general, we define **n factorial** by

$$n! = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1, \; n \geq 1$$

where n is a whole number. By convention we define

$$0! = 1$$

Example 4.1
Evaluate the following expressions: (a) $6!$ (b) $\frac{10!}{7!}$.

Solution.
(a) $6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$
(b) $\frac{10!}{7!} = \frac{10\cdot9\cdot8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1}{7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1} = 10 \cdot 9 \cdot 8 = 720$ ■

Using factorials and the Fundamental Principle of Counting, we see that the number of permutations of n objects is $n!$.

Example 4.2
There are 5! permutations of the 5 letters of the word "rehab." In how many of them is h the second letter?

Solution.
Then there are 4 ways to fill the first spot. The second spot is filled by the letter h. There are 3 ways to fill the third, 2 to fill the fourth, and one way to fill the fifth. There are 4! such permutations ■

Example 4.3
Five different books are on a shelf. In how many different ways could you arrange them?

Solution.
The five books can be arranged in $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5! = 120$ ways ■

Counting Permutations
We next consider the permutations of a set of objects taken from a larger set. Suppose we have n items. How many ordered arrangements of k items can we form from these n items? The number of permutations is denoted by ${}_nP_k$. The n refers to the number of different items and the k refers to the number of them appearing in each arrangement. A formula for ${}_nP_k$ is given next.

Theorem 4.1
For any non-negative integer n and $0 \leq k \leq n$ we have

$${}_nP_k = \frac{n!}{(n-k)!}.$$

Proof.
We can treat a permutation as a decision with k steps. The first step can be made in n different ways, the second in $n-1$ different ways, ..., the k^{th} in $n-k+1$ different ways. Thus, by the Fundamental Principle of Counting there are $n(n-1)\cdots(n-k+1)$ $k-$permutations of n objects. That is, ${}_nP_k = n(n-1)\cdots(n-k+1) = \frac{n(n-1)\cdots(n-k+1)(n-k)!}{(n-k)!} = \frac{n!}{(n-k)!}$ ■

Example 4.4
How many license plates are there that start with three letters followed by 4 digits (no repetitions)?

Solution.
The decision consists of two steps. The first is to select the letters and this can be done in ${}_{26}P_3$ ways. The second step is to select the digits and this can be done in ${}_{10}P_4$ ways. Thus, by the Fundamental Principle of Counting there are ${}_{26}P_3 \cdot {}_{10}P_4 = 78,624,000$ license plates ■

Example 4.5
How many five-digit zip codes can be made where all digits are different? The possible digits are the numbers 0 through 9.

Solution.
The answer is ${}_{10}P_5 = \frac{10!}{(10-5)!} = 30,240$ zip codes $\blacksquare$

Practice Problems

Problem 4.1
Find m and n so that ${}_mP_n = \frac{9!}{6!}$

Problem 4.2
How many four-letter code words can be formed using a standard 26-letter alphabet
(a) if repetition is allowed?
(b) if repetition is not allowed?

Problem 4.3
Certain automobile license plates consist of a sequence of three letters followed by three digits.
(a) If letters can not be repeated but digits can, how many possible license plates are there?
(b) If no letters and no digits are repeated, how many license plates are possible?

Problem 4.4
A permutation lock has 40 numbers on it.
(a) How many different three-number permutation lock can be made if the numbers can be repeated?
(b) How many different permutation locks are there if the three numbers are different?

Problem 4.5
(a) 12 cabinet officials are to be seated in a row for a picture. How many different seating arrangements are there?
(b) Seven of the cabinet members are women and 5 are men. In how many different ways can the 7 women be seated together on the left, and then the 5 men together on the right?

Problem 4.6
Using the digits 1, 3, 5, 7, and 9, with no repetitions of the digits, how many
(a) one-digit numbers can be made?
(b) two-digit numbers can be made?
(c) three-digit numbers can be made?
(d) four-digit numbers can be made?

Problem 4.7
There are five members of the Math Club. In how many ways can the positions of a president, a secretary, and a treasurer, be chosen?

Problem 4.8
Find the number of ways of choosing three initials from the alphabet if none of the letters can be repeated. Name initials such as MBF and BMF are considered different.

5 Combinations

In a permutation the order of the set of objects or people is taken into account. However, there are many problems in which we want to know the number of ways in which k objects can be selected from n distinct objects in arbitrary order. For example, when selecting a two-person committee from a club of 10 members the order in the committee is irrelevant. That is choosing Mr. A and Ms. B in a committee is the same as choosing Ms. B and Mr. A. A combination is defined as a possible selection of a certain number of objects taken from a group without regard to order. More precisely, the number of $k-$element subsets of an $n-$element set is called the number of **combinations of n objects taken k at a time**. It is denoted by ${}_nC_k$ and is read "n choose k". The formula for ${}_nC_k$ is given next.

Theorem 5.1
If ${}_nC_k$ denotes the number of ways in which k objects can be selected from a set of n distinct objects then

$${}_nC_k = \frac{{}_nP_k}{k!} = \frac{n!}{k!(n-k)!}.$$

Proof.
Since the number of groups of k elements out of n elements is ${}_nC_k$ and each group can be arranged in $k!$ ways, we have ${}_nP_k = k!{}_nC_k$. It follows that

$${}_nC_k = \frac{{}_nP_k}{k!} = \frac{n!}{k!(n-k)!} \blacksquare$$

An alternative notation for ${}_nC_k$ is $\left(\begin{array}{c} n \\ k \end{array}\right)$. We define ${}_nC_k = 0$ if $k < 0$ or $k > n$.

Example 5.1
A jury consisting of 2 women and 3 men is to be selected from a group of 5 women and 7 men. In how many different ways can this be done? Suppose that either Steve or Harry must be selected but not both, then in how many ways this jury can be formed?

Solution.
There are ${}_5C_2 \cdot {}_7C_3 = 350$ possible jury combinations consisting of 2 women

and 3 men. Now, if we suppose that Steve and Harry can not serve together then the number of jury groups that do not include the two men at the same time is ${}_5C_{25}C_{22}C_1 = 200$ ■

The next theorem discusses some of the properties of combinations.

Theorem 5.2
Suppose that n and k are whole numbers with $0 \le k \le n$. Then
(a) ${}_nC_0 =_n C_n = 1$ and ${}_nC_1 =_n C_{n-1} = n$.
(b) Symmetry property: ${}_nC_k =_n C_{n-k}$.
(c) Pascal's identity: ${}_{n+1}C_k =_n C_{k-1} +_n C_k$.

Proof.
(a) From the formula of ${}_nC_k$ we have ${}_nC_0 = \frac{n!}{0!(n-0)!} = 1$ and ${}_nC_n = \frac{n!}{n!(n-n)!} = 1$. Similarly, ${}_nC_1 = \frac{n!}{1!(n-1)!} = n$ and ${}_nC_{n-1} = \frac{n!}{(n-1)!} = n$.
(b) Indeed, we have ${}_nC_{n-k} = \frac{n!}{(n-k)!(n-n+k)!} = \frac{n!}{k!(n-k)!} =_n C_k$.
(c) We have

$$\begin{aligned}
{}_nC_{k-1} +_n C_k =& \frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!} \\
=& \frac{n!k}{k!(n-k+1)!} + \frac{n!(n-k+1)}{k!(n-k+1)!} \\
=& \frac{n!}{k!(n-k+1)!}(k+n-k+1) \\
=& \frac{(n+1)!}{k!(n+1-k)!} =_{n+1} C_k \blacksquare
\end{aligned}$$

Example 5.2
The Russellville School District has six members. In how many ways
(a) can all six members line up for a picture?
(b) can they choose a president and a secretary?
(c) can they choose three members to attend a state conference with no regard to order?

Solution.
(a) ${}_6P_6 = 6! = 720$ different ways
(b) ${}_6P_2 = 30$ ways
(c) ${}_6C_3 = 20$ different ways ■

Pascal's identity allows one to construct the so-called Pascal's triangle (for $n = 10$) as shown in Figure 5.1.

	n
1	0
1 1	1
1 2 1	2
1 3 3 1	3
1 4 6 4 1	4
1 5 10 10 5 1	5
1 6 15 20 15 6 1	6
1 7 21 35 35 21 7 1	7
1 8 28 56 70 56 28 8 1	8
1 9 36 84 126 126 84 36 9 1	9
1 10 45 120 210 252 210 120 45 10 1	10

Figure 5.1

As an application of combination we have the following theorem which provides an expansion of $(x+y)^n$, where n is a non-negative integer.

Theorem 5.3 *(Binomial Theorem)*
Let x and y be variables, and let n be a non-negative integer. Then

$$(x+y)^n = \sum_{k=0}^{n} {}_nC_k x^{n-k} y^k$$

where ${}_nC_k$ will be called the **binomial coefficient**.

Proof.
The proof is by induction on n.

Basis of induction: For $n = 0$ we have

$$(x+y)^0 = \sum_{k=0}^{0} {}_0C_k x^{0-k} y^k = 1.$$

Induction hypothesis: Suppose that the theorem is true up to n. That is,

$$(x+y)^n = \sum_{k=0}^{n} {}_nC_k x^{n-k} y^k$$

Induction step: Let us show that it is still true for $n+1$. That is

$$(x+y)^{n+1} = \sum_{k=0}^{n+1} {}_{n+1}C_k x^{n-k+1} y^k.$$

Indeed, we have

$$\begin{aligned}
(x+y)^{n+1} =& (x+y)(x+y)^n = x(x+y)^n + y(x+y)^n \\
=& x\sum_{k=0}^{n} {}_nC_k x^{n-k}y^k + y\sum_{k=0}^{n} {}_nC_k x^{n-k}y^k \\
=& \sum_{k=0}^{n} {}_nC_k x^{n-k+1}y^k + \sum_{k=0}^{n} {}_nC_k x^{n-k}y^{k+1} \\
=& [{}_nC_0x^{n+1} + {}_nC_1x^ny + {}_nC_2x^{n-1}y^2 + \cdots + {}_nC_nxy^n] \\
+& [{}_nC_0x^ny + {}_nC_1x^{n-1}y^2 + \cdots + {}_nC_{n-1}xy^n + {}_nC_ny^{n+1}] \\
=& {}_{n+1}C_0x^{n+1} + [{}_nC_1 + {}_nC_0]x^ny + \cdots + \\
& [{}_nC_n + {}_nC_{n-1}]xy^n + {}_{n+1}C_{n+1}y^{n+1} \\
=& {}_{n+1}C_0x^{n+1} + {}_{n+1}C_1x^ny + {}_{n+1}C_2x^{n-1}y^2 + \cdots \\
+& {}_{n+1}C_nxy^n + {}_{n+1}C_{n+1}y^{n+1} \\
=& \sum_{k=0}^{n+1} {}_{n+1}C_k x^{n-k+1}y^k. \blacksquare
\end{aligned}$$

Note that the coefficients in the expansion of $(x+y)^n$ are the entries of the $(n+1)^{\text{st}}$ row of Pascal's triangle.

Example 5.3
Expand $(x+y)^6$ using the Binomial Theorem.

Solution.
By the Binomial Theorem and Pascal's triangle we have

$$(x+y)^6 = x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^4 + 6xy^5 + y^6 \blacksquare$$

Example 5.4
How many subsets are there of a set with n elements?

Solution.
Since there are ${}_nC_k$ subsets of k elements with $0 \leq k \leq n$, the total number of subsets of a set of n elements is

$$\sum_{k=0}^{n} {}_nC_k = (1+1)^n = 2^n \blacksquare$$

Practice Problems

Problem 5.1
Find m and n so that ${}_mC_n = 13$

Problem 5.2
A club with 42 members has to select three representatives for a regional meeting. How many possible choices are there?

Problem 5.3
In a UN ceremony, 25 diplomats were introduced to each other. Suppose that the diplomats shook hands with each other exactly once. How many handshakes took place?

Problem 5.4
There are five members of the math club. In how many ways can the two-person Social Committee be chosen?

Problem 5.5
A medical research group plans to select 2 volunteers out of 8 for a drug experiment. In how many ways can they choose the 2 volunteers?

Problem 5.6
A consumer group has 30 members. In how many ways can the group choose 3 members to attend a national meeting?

Problem 5.7
Which is usually greater the number of combinations of a set of objects or the number of permutations?

Problem 5.8
Determine whether each problem requires a combination or a permutation:
(a) There are 10 toppings available for your ice cream and you are allowed to choose only three. How many possible 3-topping combinations can yo have?
(b) Fifteen students participated in a spelling bee competition. The first place winner will receive $1,000, the second place $500, and the third place $250. In how many ways can the 3 winners be drawn?

Problem 5.9
Use the binomial theorem and Pascal's triangle to find the expansion of $(a+b)^7$.

Problem 5.10
Find the 5^{th} term in the expansion of $(2a-3b)^7$.

Problem 5.11 ‡
Thirty items are arranged in a 6-by-5 array as shown.

A_1	A_2	A_3	A_4	A_5
A_6	A_7	A_8	A_9	A_{10}
A_{11}	A_{12}	A_{13}	A_{14}	A_{15}
A_{16}	A_{17}	A_{18}	A_{19}	A_{20}
A_{21}	A_{22}	A_{23}	A_{24}	A_{25}
A_{26}	A_{27}	A_{28}	A_{29}	A_{30}

Calculate the number of ways to form a set of three distinct items such that no two of the selected items are in the same row or same column.

Probability: Definitions and Properties

In this chapter we discuss the fundamental concepts of probability at a level at which no previous exposure to the topic is assumed.
Probability has been used in many applications ranging from medicine to business and so the study of probability is considered an essential component of any mathematics curriculum.
So what is probability? Before answering this question we start with some basic definitions.

6 Sample Space, Events, Probability Measure

A **random experiment** or simply an **experiment** is an experiment whose **outcomes** cannot be predicted with certainty. Examples of an experiment include rolling a die, flipping a coin, and choosing a card from a deck of playing cards.
The **sample space** S of an experiment is the set of all possible outcomes for the experiment. For example, if you roll a die one time then the experiment is the roll of the die. A sample space for this experiment could be $S = \{1, 2, 3, 4, 5, 6\}$ where each digit represents a face of the die.
An **event** is a subset of the sample space. For example, the event of rolling an odd number with a die consists of three outcomes $\{1, 3, 5\}$.

Example 6.1
Consider the random experiment of tossing a coin three times.
(a) Find the sample space of this experiment.
(b) Find the outcomes of the event of obtaining more than one head.

Solution.
We will use T for tail and H for head.
(a) The sample space is composed of eight outcomes:

$$S = \{TTT, TTH, THT, THH, HTT, HTH, HHT, HHH\}.$$

(b) The event of obtaining more than one head is the set

$$\{THH, HTH, HHT, HHH\} \blacksquare$$

Probability is the measure of occurrence of an event. Various probability concepts exist nowadays. A widely used probability concept is the **experimental** probability which uses the relative frequency of an event and is defined as follows. Let $n(E)$ denote the number of times in the first n repetitions of the experiment that the event E occurs. Then $\Pr(E)$, the probability of the event E, is defined by

$$\Pr(E) = \lim_{n\to\infty} \frac{n(E)}{n}.$$

This states that if we repeat an experiment a large number of times then the fraction of times the event E occurs will be close to $\Pr(E)$. This result is a theorem called the law of large numbers which we will discuss in Section 50.1.
The function Pr satisfies the following axioms, known as **Kolmogorov axioms**:
Axiom 1: For any event E, $0 \le \Pr(E) \le 1$.
Axiom 2: $\Pr(S) = 1$.
Axiom 3: For any sequence of mutually exclusive events $\{E_n\}_{n\ge 1}$, that is $E_i \cap E_j = \emptyset$ for $i \neq j$, we have

$$\Pr\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \Pr(E_n).\text{(Countable additivity)}$$

If we let $E_1 = S$, $E_n = \emptyset$ for $n > 1$ then by Axioms 2 and 3 we have $1 = \Pr(S) = \Pr\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \Pr(E_n) = \Pr(S) + \sum_{n=2}^{\infty} \Pr(\emptyset)$. This implies that $\Pr(\emptyset) = 0$. Also, if $\{E_1, E_2, \cdots, E_n\}$ is a finite set of mutually exclusive events, then by defining $E_k = \emptyset$ for $k > n$ and Axioms 3 we find

$$\Pr\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \Pr(E_k).$$

Any function Pr that satisfies Axioms 1 - 3 will be called a **probability measure**.

Example 6.2
Consider the sample space $S = \{1, 2, 3\}$. Suppose that $\Pr(\{1, 2\}) = 0.5$ and $\Pr(\{2, 3\}) = 0.7$. Is Pr a valid probability measure? Justify your answer.

Solution.
We have $\Pr(\{1\}) + \Pr(\{2\}) + \Pr(\{3\}) = 1$. But $\Pr(\{1, 2\}) = \Pr(\{1\}) + \Pr(\{2\}) = 0.5$. This implies that $0.5 + \Pr(\{3\}) = 1$ or $\Pr(\{3\}) = 0.5$. Similarly, $1 = \Pr(\{2, 3\}) + \Pr(\{1\}) = 0.7 + \Pr(\{1\})$ and so $\Pr(\{1\}) = 0.3$. It follows that $\Pr(\{2\}) = 1 - \Pr(\{1\}) - \Pr(\{3\}) = 1 - 0.3 - 0.5 = 0.2$. Since $\Pr(\{1\}) + \Pr(\{2\}) + \Pr(\{3\}) = 1$, Pr is a valid probability measure ■

Example 6.3
If, for a given experiment, $O_1, O_2, O_3, \cdots$ is an infinite sequence of distinct outcomes, verify that

$$\Pr(\{O_i\}) = \left(\frac{1}{2}\right)^i, \; i = 1, 2, 3, \cdots$$

is a probability measure.

Solution.
Note that $\Pr(E) > 0$ for any event E. Moreover, if S is the sample space then

$$\Pr(S) = \sum_{i=1}^{\infty} \Pr(\{O_i\}) = \frac{1}{2} \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^i = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{2}} = 1$$

where the infinite sum is the infinite geometric series

$$1 + a + a^2 + \cdots + a^n + \cdots = \frac{1}{1 - a}, |a| < 1$$

with $a = \frac{1}{2}$.
Next, if $E_1, E_2, \cdots$ is a sequence of mutually exclusive events then

$$\Pr\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} \Pr(\{O_{nj}\}) = \sum_{n=1}^{\infty} \Pr(E_n)$$

where $E_n = \cup_{i=1}^{\infty}\{O_{ni}\}$. Thus, Pr defines a probability function ■

Now, since $E \cup E^c = S,\ E \cap E^c = \emptyset$, and $\Pr(S) = 1$ we find

$$\Pr(E^c) = 1 - \Pr(E)$$

where E^c is the **complementary event**.
When the outcome of an experiment is just as likely as another, as in the example of tossing a coin, the outcomes are said to be **equally likely**. The **classical probability** concept applies *only* when all possible outcomes are equally likely, in which case we use the formula

$$\Pr(E) = \frac{\text{number of outcomes favorable to event}}{\text{total number of outcomes}} = \frac{n(E)}{n(S)}.$$

Since for any event E we have $\emptyset \subseteq E \subseteq S$, we can write $0 \le n(E) \le n(S)$ so that $0 \le \frac{n(E)}{n(S)} \le 1$. It follows that $0 \le \Pr(E) \le 1$. Clearly, $\Pr(S) = 1$. Also, Axiom 3 is easy to check using a generalization of Theorem 2.3 (b).

Example 6.4
A hand of 5 cards is dealt from a deck. Let E be the event that the hand contains 5 aces. List the elements of E and find $\Pr(E)$.

Solution.
Recall that a standard deck of 52 playing cards can be described as follows:

hearts (red)	Ace	2	3	4	5	6	7	8	9	10	Jack	Queen	King
clubs (black)	Ace	2	3	4	5	6	7	8	9	10	Jack	Queen	King
diamonds (red)	Ace	2	3	4	5	6	7	8	9	10	Jack	Queen	King
spades (black)	Ace	2	3	4	5	6	7	8	9	10	Jack	Queen	King

Cards labeled Ace, Jack, Queen, or King are called **face cards.**
Since there are only 4 aces in the deck, event E is impossible, i.e. $E = \emptyset$ so that $\Pr(E) = 0$ ■

Example 6.5
What is the probability of drawing an ace from a well-shuffled deck of 52 playing cards?

Solution.
Since there are four aces in a deck of 52 playing cards, the probability of getting an ace is $\frac{4}{52} = \frac{1}{13}$ ■

Example 6.6
What is the probability of rolling a 3 or a 4 with a fair die?

Solution.
The event of having a 3 or a 4 has two outcomes $\{3, 4\}$. The probability of rolling a 3 or a 4 is $\frac{2}{6} = \frac{1}{3}$ ■

Example 6.7 *(Birthday problem)*
In a room containing n people, calculate the chance that at least two of them have the same birthday.

Solution.
In a group of n randomly chosen people, the sample space S will consist of all ordered $n-$tuples of birthdays. Let S_i be denote the birthday of the i^{th} person, where $1 \leq i \leq n$. Then $n(S_i) = 365$ (assuming no leap year). Moreover, $S = S_1 \times S_2 \times \cdots \times S_n$. Hence,

$$n(S) = n(S_1)n(S_2)\cdots n(S_n) = 365^n.$$

Now, let E be the event that at least two people share the same birthday. Then the complementary event E^c is the event that no two people of the n people share the same birthday. Moreover,

$$\Pr(E) = 1 - \Pr(E^c).$$

The outcomes in E^c are ordered arrangements of n numbers chosen from 365 numbers without repetitions. Therefore

$$n(E^c) = {}_{365}P_n = (365)(364)\cdots(365 - n + 1).$$

Hence,

$$\Pr(E^c) = \tfrac{(365)(364)\cdots(365-n+1)}{(365)^n}$$

and

$$\Pr(E) = 1 - \frac{(365)(364)\cdots(365 - n + 1)}{(365)^n} \blacksquare$$

Remark 6.1
It is important to keep in mind that the classical definition of probability applies only to a sample space that has equally likely outcomes. Applying the definition to a space with outcomes that are not equally likely leads to incorrect conclusions. For example, the sample space for spinning the spinner in Figure 6.1 is given by $S = \{Red, Blue\}$, but the outcome Blue is more likely to occur than is the outcome Red. Indeed, $\Pr(Blue) = \frac{3}{4}$ whereas $\Pr(Red) = \frac{1}{4}$ as opposed to $\Pr(Blue) = \Pr(Red) = \frac{1}{2}$

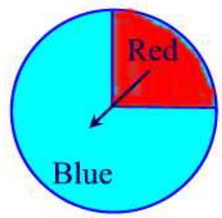


Figure 6.1

Practice Problems

Problem 6.1
Consider the random experiment of rolling a die.
(a) Find the sample space of this experiment.
(b) Find the event of rolling the die an even number.

Problem 6.2
An experiment consists of the following two stages: (1) first a coin is tossed (2) if the face appearing is a head, then a die is rolled; if the face appearing is a tail, then the coin is tossed again. An outcome of this experiment is a pair of the form (outcome from stage 1, outcome from stage 2). Let S be the collection of all outcomes.
Find the sample space of this experiment.

Problem 6.3 ‡
An insurer offers a health plan to the employees of a large company. As part of this plan, the individual employees may choose exactly two of the supplementary coverages A, B, and C, or they may choose no supplementary coverage. The proportions of the company's employees that choose coverages A, B, and C are $\frac{1}{4}, \frac{1}{3}$, and , $\frac{5}{12}$ respectively.
Determine the probability that a randomly chosen employee will choose no supplementary coverage.

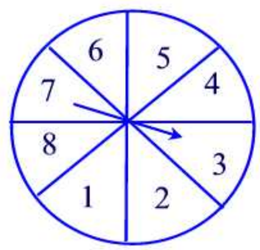
Problem 6.4
An experiment consists of throwing two dice.
(a) Write down the sample space of this experiment.
(b) If E is the event "total score is at most 10", list the outcomes belonging to E^c.
(c) Find the probability that the total score is at most 10 when the two dice are thrown.
(d) What is the probability that a double, that is,

$$\{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$$

will not be thrown?
(e) What is the probability that a double is not thrown nor is the score greater than 10?

Problem 6.5
Let $S = \{1, 2, 3, \cdots, 10\}$. If a number is chosen at **random**, that is, with the same chance of being drawn as all other numbers in the set, calculate each of the following probabilities:
(a) The event A that an even number is drawn.
(b) The event B that a number less than 5 and greater than 9 is drawn.
(c) The event C that a number less than 11 but greater than 0 is drawn.
(d) The event D that a prime number is drawn.
(e) The event E that a number both odd and prime is drawn.

Problem 6.6
The following spinner is spun:



Find the probabilities of obtaining each of the following:
(a) Pr(factor of 24)
(b) Pr(multiple of 4)
(c) Pr(odd number)
(d) Pr($\{9\}$)
(e) Pr(composite number), i.e., a number that is not prime
(f) Pr(neither prime nor composite)

Problem 6.7
A box of clothes contains 15 shirts and 10 pants. Three items are drawn from the box without replacement. What is the probability that all three are all shirts or all pants?

Problem 6.8
A coin is tossed repeatedly. What is the probability that the second head appears at the 7^{th} toss? (Hint: Since only the first seven tosses matter, you can assume that the coin is tossed only 7 times.)

Problem 6.9
Suppose each of 100 professors in a large mathematics department picks at random one of 200 courses. What is the probability that at least two professors pick the same course?

Problem 6.10
A large classroom has 100 foreign students, 30 of whom speak spanish. 25 of the students speak italian, while 55 do not speak neither spanish nor italian.
(a) How many of the those speak both spanish and italian?
(b) A student who speak italian is chosen at random. What is the probability that he/she speaks spanish?

Problem 6.11
A box contains 5 batteries of which 2 are defective. An inspector selects 2 batteries at random from the box. She/he tests the 2 items and observes whether the sampled items are defective.
(a) Write out the sample space of all possible outcomes of this experiment. Be very specific when identifying these.
(b) The box will not be accepted if both of the sampled items are defective. What is the probability the inspector will reject the box?

7 Probability of Intersection, Union, and Complementary Event

In this section we find the probability of a complementary event, the union of two events and the intersection of two events.
We define the probability of nonoccurrence of an event E (called its **failure** or the **complementary event**) to be the number $\Pr(E^c)$. Since $S = E \cup E^c$ and $E \cap E^c = \emptyset$, $\Pr(S) = \Pr(E) + \Pr(E^c)$. Thus,

$$\Pr(E^c) = 1 - \Pr(E).$$

Example 7.1
The probability that a senior citizen in a nursing home without a pneumonia shot will get pneumonia is 0.45. What is the probability that a senior citizen without pneumonia shot will not get pneumonia?

Solution.
Our sample space consists of those senior citizens in the nursing home who did not get the pneumonia shot. Let E be the set of those individuals without the shot who did get the illness. Then $\Pr(E) = 0.45$. The probability that an individual without the shot will not get the illness is then $\Pr(E^c) = 1 - \Pr(E) = 1 - 0.45 = 0.55$ ■

The **union** of two events A and B is the event $A \cup B$ whose outcomes are either in A or in B. The **intersection** of two events A and B is the event $A \cap B$ whose outcomes are outcomes of both events A and B. Two events A and B are said to be **mutually exclusive** if they have no outcomes in common. In this case $A \cap B = \emptyset$ and $\Pr(A \cap B) = \Pr(\emptyset) = 0$.

Example 7.2
Consider the sample space of rolling a die. Let A be the event of rolling a prime number, B the event of rolling a composite number, and C the event of rolling a 4. Find
(a) $A \cup B, A \cup C$, and $B \cup C$.
(b) $A \cap B, A \cap C$, and $B \cap C$.
(c) Which events are mutually exclusive?

Solution.
(a) We have

$$\begin{aligned}A \cup B &= \{2, 3, 4, 5, 6\}\\ A \cup C &= \{2, 3, 4, 5\}\\ B \cup C &= \{4, 6\}\end{aligned}$$

(b)

$$\begin{aligned}A \cap B &= \emptyset\\ A \cap C &= \emptyset\\ B \cap C &= \{4\}\end{aligned}$$

(c) A and B are mutually exclusive as well as A and C ■

Example 7.3
Let A be the event of drawing a "Queen" from a well-shuffled standard deck of playing cards and B the event of drawing an "ace" card. Are A and B mutually exclusive?

Solution.
Since $A = \{\text{queen of diamonds, queen of hearts, queen of clubs, queen of spades}\}$ and $B = \{\text{ace of diamonds, ace of hearts, ace of clubs, ace of spades}\}$, A and B are mutually exclusive ■

For any events A and B the probability of $A \cup B$ is given by the addition rule.

Theorem 7.1
Let A and B be two events. Then

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

Proof.
Let $A^c \cap B$ denote the event whose outcomes are the outcomes in B that are not in A. Then using the Venn diagram in Figure 7.1 we see that $B = (A \cap B) \cup (A^c \cap B)$ and $A \cup B = A \cup (A^c \cap B)$.

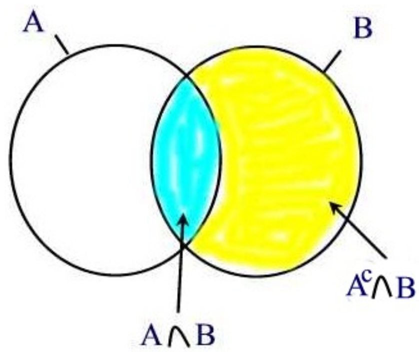


Figure 7.1

Since $(A \cap B)$ and $(A^c \cap B)$ are mutually exclusive, by Axiom 3 of Section 6, we have

$$\Pr(B) = \Pr(A \cap B) + \Pr(A^c \cap B).$$

Thus,

$$\Pr(A^c \cap B) = \Pr(B) - \Pr(A \cap B).$$

Similarly, A and $A^c \cap B$ are mutually exclusive, thus we have

$$\Pr(A \cup B) = \Pr(A) + \Pr(A^c \cap B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \blacksquare$$

Note that in the case A and B are mutually exclusive, $\Pr(A \cap B) = 0$ so that

$$\Pr(A \cup B) = \Pr(A) + \Pr(B).$$

Example 7.4
An airport security has two checkpoints. Let A be the event that the first checkpoint is busy, and let B be the event the second checkpoint is busy. Assume that $\Pr(A) = 0.2$, $\Pr(B) = 0.3$ and $\Pr(A \cap B) = 0.06$. Find the probability that neither of the two checkpoints is busy.

Solution.
The probability that neither of the checkpoints is busy is $\Pr[(A \cup B)^c] = 1 - \Pr(A \cup B)$. But $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = 0.2 + 0.3 - 0.06 = 0.44$. Hence, $\Pr[(A \cup B)^c] = 1 - 0.44 = 0.56$ $\blacksquare$

Example 7.5
Let $\Pr(A) = 0.9$ and $\Pr(B) = 0.6$. Find the minimum possible value for $\Pr(A \cap B)$.

Solution.
Since $\Pr(A) + \Pr(B) = 1.5$ and $0 \le \Pr(A \cup B) \le 1$, by the previous theorem

$$\Pr(A \cap B) = \Pr(A) + \Pr(B) - \Pr(A \cup B) \ge 1.5 - 1 = 0.5.$$

So the minimum value of $\Pr(A \cap B)$ is 0.5 ■

Example 7.6
Suppose there's 40% chance of getting a freezing rain, 10% chance of snow and freezing rain, 80% chance of snow or freezing rain. Find the chance of snow.

Solution.
By the addition rule we have

$$\Pr(R) = \Pr(R \cup C) - \Pr(C) + \Pr(R \cap C) = 0.8 - 0.4 + 0.1 = 0.5 \blacksquare$$

Example 7.7
Let $\mathbb{N}$ be the set of all positive integers and Pr be a probability measure defined by $\Pr(n) = \left(\frac{1}{3}\right)^n$ for all $n \in \mathbb{N}$. What is the probability that a number chosen at random from $\mathbb{N}$ will be odd?

Solution.
We have

$$\begin{aligned}
\Pr(\{1,3,5,\cdots\}) =& \Pr(\{1\}) + \Pr(\{3\}) + \Pr(\{5\}) + \cdots \\
=& \left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)^3 + \left(\frac{1}{3}\right)^5 + \cdots \\
=& \left(\frac{1}{3}\right)\left[1 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^4 + \cdots\right] \\
=& \left(\frac{1}{3}\right) \cdot \frac{1}{1 - \left(\frac{1}{3}\right)^2} = \frac{3}{8} \blacksquare
\end{aligned}$$

Finally, if E and F are two events such that $E \subseteq F$, then F can be written as the union of two mutually exclusive events $F = E \cup (E^c \cap F)$. By Axiom 3 we obtain

$$\Pr(F) = \Pr(E) + \Pr(E^c \cap F).$$

Thus, $\Pr(F) - \Pr(E) = \Pr(E^c \cap F) \ge 0$ and this shows

$$E \subseteq F \Longrightarrow \Pr(E) \le \Pr(F).$$

Theorem 7.2
For any three events A, B, and C we have

$$\begin{aligned}\Pr(A\cup B\cup C) =&\Pr(A)+\Pr(B)+\Pr(C)-\Pr(A\cap B)-\Pr(A\cap C)-\Pr(B\cap C)\\&+\Pr(A\cap B\cap C).\end{aligned}$$

Proof.
We have

$$\begin{aligned}\Pr(A\cup B\cup C) =& \Pr(A)+\Pr(B\cup C)-\Pr(A\cap(B\cup C))\\=&\Pr(A)+\Pr(B)+\Pr(C)-\Pr(B\cap C)\\&-\Pr((A\cap B)\cup(A\cap C))\\=&\Pr(A)+\Pr(B)+\Pr(C)-\Pr(B\cap C)\\&-[\Pr(A\cap B)+\Pr(A\cap C)-\Pr((A\cap B)\cap(A\cap C))]\\=&\Pr(A)+\Pr(B)+\Pr(C)-\Pr(B\cap C)\\&-\Pr(A\cap B)-\Pr(A\cap C)+\Pr((A\cap B)\cap(A\cap C))\\=&\Pr(A)+\Pr(B)+\Pr(C)-\Pr(A\cap B)-\\&\Pr(A\cap C)-\Pr(B\cap C)+\Pr(A\cap B\cap C)\ \blacksquare\end{aligned}$$

Example 7.8
If a person visits his primary care physician, suppose that the probability that he will have blood test work is 0.44, the probability that he will have an X-ray is 0.24, the probability that he will have an MRI is 0.21, the probability that he will have blood test and an X-ray is 0.08, the probability that he will have blood test and an MRI is 0.11, the probability that he will have an X-ray and an MRI is 0.07, and the probability that he will have blood test, an X-ray, and an MRI is 0.03. What is the probability that a person visiting his PCP will have at least one of these things done to him/her?

Solution.
Let B be the event that a person will have blood test, X is the event that a person will have an X-ray, and M is the event a person will have an MRI. We are given $\Pr(B) = 0.44, \Pr(X) = 0.24, \Pr(M) = 0.21, \Pr(B \cap X) = 0.08, \Pr(B\cap M) = 0.11, \Pr(X\cap M) = 0.07$ and $\Pr(B\cap X\cap M) = 0.03$. Thus,

$$\Pr(B\cup X\cup M) = 0.44 + 0.24 + 0.21 - 0.08 - 0.11 - 0.07 + 0.03 = 0.66\ \blacksquare$$

Practice Problems

Problem 7.1
A consumer testing service rates a given DVD player as either very good or good. Let A denote the event that the rating is very good and B the event that the rating is good. You are given: $\Pr(A) = 0.22$, $\Pr(B) = 0.35$. Find
(a) $\Pr(A^c)$;
(b) $\Pr(A \cup B)$;
(c) $\Pr(A \cap B)$.

Problem 7.2
An entrance exam consists of two subjects: Math and english. The probability that a student fails the math test is 0.20. The probability of failing english is 0.15, and the probability of failing both subjects is 0.03. What is the probability that the student will fail at least one of these subjects?

Problem 7.3
Let A be the event of "drawing a king" from a deck of cards and B the event of "drawing a diamond". Are A and B mutually exclusive? Find $\Pr(A \cup B)$.

Problem 7.4
An urn contains 4 red balls, 8 yellow balls, and 6 green balls. A ball is selected at random. What is the probability that the ball chosen is either red or green?

Problem 7.5
Show that for any events A and B, $\Pr(A \cap B) \geq \Pr(A) + \Pr(B) - 1$.

Problem 7.6
An urn contains 2 red balls, 4 blue balls, and 5 white balls.
(a) What is the probability of the event R that a ball drawn at random is red?
(b) What is the probability of the event "not R" that is, that a ball drawn at random is not red?
(c) What is the probability of the event that a ball drawn at random is either red or blue?

Problem 7.7
In the experiment of rolling of fair pair of dice, let E denote the event of rolling a sum that is an even number and P the event of rolling a sum that is a prime number. Find the probability of rolling a sum that is even or prime?

Problem 7.8
Let S be a sample space and A and B be two events such that $\Pr(A) = 0.8$ and $\Pr(B) = 0.9$. Determine whether A and B are mutually exclusive or not.

Problem 7.9 ‡
A survey of a group's viewing habits over the last year revealed the following information

(i) 28% watched gymnastics
(ii) 29% watched baseball
(iii) 19% watched soccer
(iv) 14% watched gymnastics and baseball
(v) 12% watched baseball and soccer
(vi) 10% watched gymnastics and soccer
(vii) 8% watched all three sports.

Find the probability of the group that watched none of the three sports during the last year.

Problem 7.10 ‡
The probability that a visit to a primary care physician's (PCP) office results in neither lab work nor referral to a specialist is 35% . Of those coming to a PCP's office, 30% are referred to specialists and 40% require lab work.
Determine the probability that a visit to a PCP's office results in both lab work and referral to a specialist.

Problem 7.11 ‡
You are given $\Pr(A \cup B) = 0.7$ and $\Pr(A \cup B^c) = 0.9$. Determine $\Pr(A)$.

Problem 7.12 ‡
Among a large group of patients recovering from shoulder injuries, it is found that 22% visit both a physical therapist and a chiropractor, whereas 12% visit neither of these. The probability that a patient visits a chiropractor exceeds by 14% the probability that a patient visits a physical therapist.
Determine the probability that a randomly chosen member of this group visits a physical therapist.

Problem 7.13 ‡
In modeling the number of claims filed by an individual under an automobile policy during a three-year period, an actuary makes the simplifying

assumption that for all integers $n \geq 0, p_{n+1} = \frac{1}{5}p_n$, where p_n represents the probability that the policyholder files n claims during the period.
Under this assumption, what is the probability that a policyholder files more than one claim during the period?

Problem 7.14 ‡
A marketing survey indicates that 60% of the population owns an automobile, 30% owns a house, and 20% owns both an automobile and a house.
Calculate the probability that a person chosen at random owns an automobile or a house, but not both.

Problem 7.15 ‡
An insurance agent offers his clients auto insurance, homeowners insurance and renters insurance. The purchase of homeowners insurance and the purchase of renters insurance are mutually exclusive. The profile of the agent's clients is as follows:
i) 17% of the clients have none of these three products.
ii) 64% of the clients have auto insurance.
iii) Twice as many of the clients have homeowners insurance as have renters insurance.
iv) 35% of the clients have two of these three products.
v) 11% of the clients have homeowners insurance, but not auto insurance.
Calculate the percentage of the agent's clients that have both auto and renters insurance.

Problem 7.16 ‡
A mattress store sells only king, queen and twin-size mattresses. Sales records at the store indicate that one-fourth as many queen-size mattresses are sold as king and twin-size mattresses combined. Records also indicate that three times as many king-size mattresses are sold as twin-size mattresses.
Calculate the probability that the next mattress sold is either king or queen-size.

Problem 7.17 ‡
The probability that a member of a certain class of homeowners with liability and property coverage will file a liability claim is 0.04, and the probability that a member of this class will file a property claim is 0.10. The probability that a member of this class will file a liability claim but not a property claim

is 0.01.
Calculate the probability that a randomly selected member of this class of homeowners will not file a claim of either type.

8 Probability and Counting Techniques

The Fundamental Principle of Counting can be used to compute probabilities as shown in the following example.

Example 8.1
In an actuarial course in probability, an instructor has decided to give his class a weekly quiz consisting of 5 multiple-choice questions taken from a pool of previous SOA P/1 exams. Each question has 4 answer choices, of which 1 is correct and the other 3 are incorrect.
(a) How many answer choices are there?
(b) What is the probability of getting all 5 right answers?
(c) What is the probability of answering exactly 4 questions correctly?
(d) What is the probability of getting at least four answers correctly?

Solution.
(a) We can look at this question as a decision consisting of five steps. There are 4 ways to do each step so that by the Fundamental Principle of Counting there are

$$(4)(4)(4)(4)(4) = 1024 \text{ possible choices of answers.}$$

(b) There is only one way to answer each question correctly. Using the Fundamental Principle of Counting there is $(1)(1)(1)(1)(1) = 1$ way to answer all 5 questions correctly out of 1024 possible answer choices. Hence,

$$\Pr(\text{all 5 right}) = \tfrac{1}{1024}$$

(c) The following table lists all possible responses that involve exactly 4 right answers where R stands for right and W stands for a wrong answer

Five Responses	Number of ways to fill out the test
WRRRR	$(3)(1)(1)(1)(1) = 3$
RWRRR	$(1)(3)(1)(1)(1) = 3$
RRWRR	$(1)(1)(3)(1)(1) = 3$
RRRWR	$(1)(1)(1)(3)(1) = 3$
RRRRW	$(1)(1)(1)(1)(3) = 3$

So there are 15 ways out of the 1024 possible ways that result in 4 right answers and 1 wrong answer so that

$$\Pr(4 \text{ right,1 wrong}) = \tfrac{15}{1024} \approx 1.5\%$$

(d) "At least 4" means you can get either 4 right and 1 wrong or all 5 right. Thus,

$$\begin{aligned}\Pr(\text{at least 4 right}) =&\Pr(4R, 1W) + P(5R)\\ =&\frac{15}{1024} + \frac{1}{1024}\\ =&\frac{16}{1024} \approx 0.016 \blacksquare\end{aligned}$$

Example 8.2
Consider the experiment of rolling two dice. How many events A are there with $\Pr(A) = \frac{1}{3}$?

Solution.
We must have $\Pr(\{i, j\}) = \frac{1}{3}$ with $i \neq j$. There are ${}_6C_2 = 15$ such events $\blacksquare$

Probability Trees
Probability trees can be used to compute the probabilities of combined outcomes in a sequence of experiments.

Example 8.3
Construct the probability tree of the experiment of flipping a fair coin twice.

Solution.
The probability tree is shown in Figure 8.1$\blacksquare$

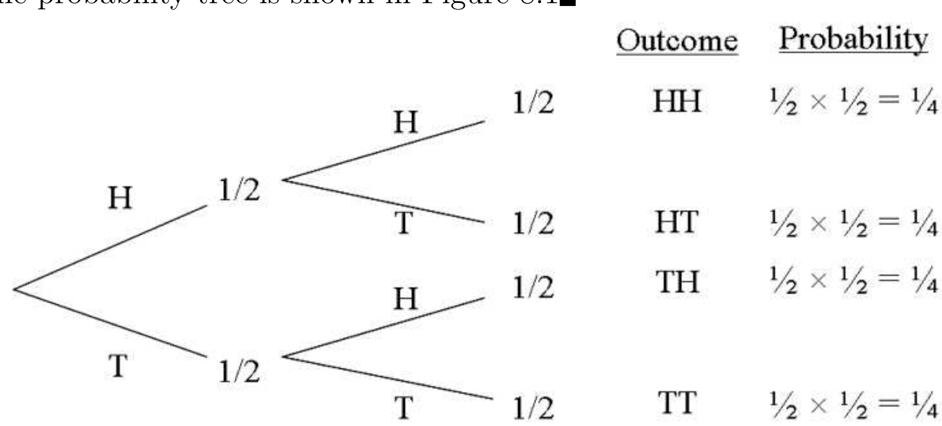


Figure 8.1

The probabilities shown in Figure 8.1 are obtained by following the paths leading to each of the four outcomes and multiplying the probabilities along the paths. This procedure is an instance of the following general property.

Multiplication Rule for Probabilities for Tree Diagrams
For all multistage experiments, the probability of the outcome along any path of a tree diagram is equal to the product of all the probabilities along the path.

Example 8.4
A shipment of 500 DVD players contains 9 defective DVD players. Construct the probability tree of the experiment of sampling two of them without replacement.

Solution.
The probability tree is shown in Figure 8.2 ■

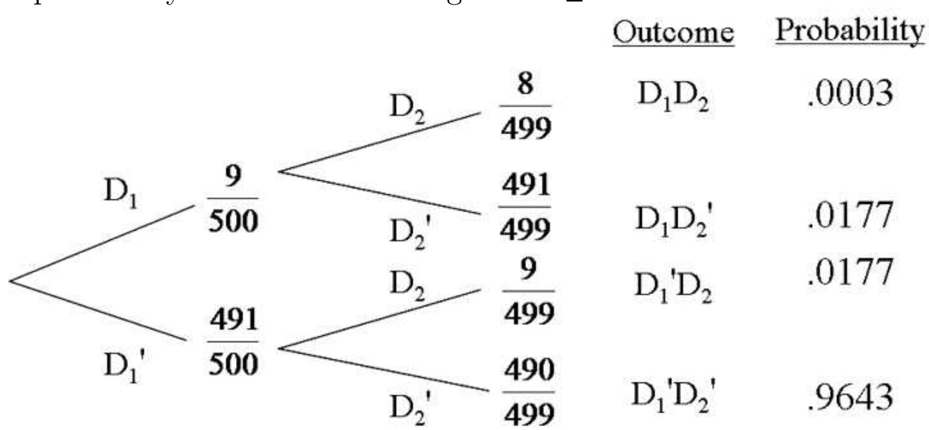


Figure 8.2

Example 8.5
The faculty of a college consists of 35 female faculty and 65 male faculty. 70% of the female faculty favor raising tuition, while only 40% of the male faculty favor the increase.
If a faculty member is selected at random from this group, what is the probability that he or she favors raising tuition?

Solution.
Figure 8.3 shows a tree diagram for this problem where F stands for female,

M for male.

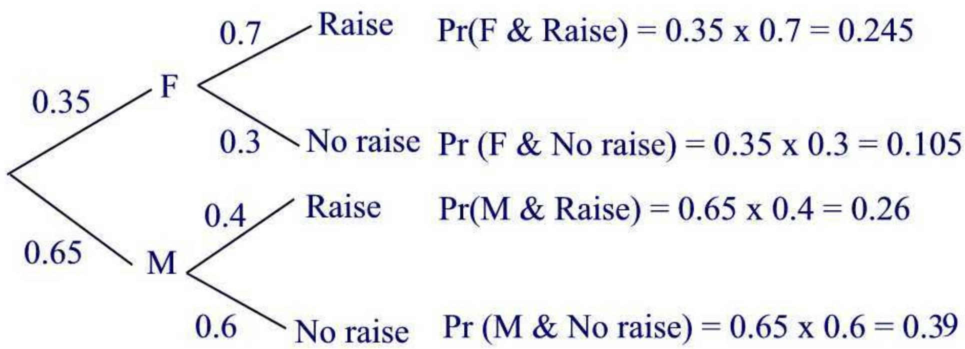


Figure 8.3

The first and third branches correspond to favoring the tuition raise. We add their probabilities.

$$\Pr(\text{tuition raise}) = 0.245 + 0.26 = 0.505 \blacksquare$$

Example 8.6

A regular insurance claimant is trying to hide 3 fraudulent claims among 7 genuine claims. The claimant knows that the insurance company processes claims in batches of 5 or in batches of 10. For batches of 5, the insurance company will investigate one claim at random to check for fraud; for batches of 10, two of the claims are randomly selected for investigation. The claimant has three possible strategies:
(a) submit all 10 claims in a single batch,
(b) submit two batches of 5, one containing 2 fraudulent claims and the other containing 1,
(c) submit two batches of 5, one containing 3 fraudulent claims and the other containing 0.
What is the probability that all three fraudulent claims will go undetected in each case? What is the best strategy?

Solution.
(a) $\Pr(\text{fraud not detected}) = \frac{7}{10} \cdot \frac{6}{9} = \frac{7}{15}$
(b) $\Pr(\text{fraud not detected}) = \frac{3}{5} \cdot \frac{4}{5} = \frac{12}{25}$
(c) $\Pr(\text{fraud not detected}) = \frac{2}{5} \cdot 1 = \frac{2}{5}$
Claimant's best strategy is to split fraudulent claims between two batches of 5 $\blacksquare$

Practice Problems

Problem 8.1
A box contains three red balls and two blue balls. Two balls are to be drawn without replacement. Use a tree diagram to represent the various outcomes that can occur. What is the probability of each outcome?

Problem 8.2
Repeat the previous exercise but this time replace the first ball before drawing the second.

Problem 8.3
An urn contains three red marbles and two green marbles. An experiment consists of drawing one marble at a time without replacement, until a red one is obtained. Find the probability of the following events.

A : Only one draw is needed.
B : Exactly two draws are needed.
C : Exactly three draws are needed.

Problem 8.4
Consider a jar with three black marbles and one red marble. For the experiment of drawing two marbles with replacement, what is the probability of drawing a black marble and then a red marble in that order? Assume that the balls are equally likely to be drawn.

Problem 8.5
An urn contains two black balls and one red ball. Two balls are drawn with replacement. What is the probability that both balls are black? Assume that the balls are equally likely to be drawn.

Problem 8.6
An urn contains four balls: one red, one green, one yellow, and one white. Two balls are drawn without replacement from the urn. What is the probability of getting a red ball and a white ball? Assume that the balls are equally likely to be drawn

Problem 8.7
An urn contains 3 white balls and 2 red balls. Two balls are to be drawn one at a time and without replacement. Draw a tree diagram for this experiment and find the probability that the two drawn balls are of different colors. Assume that the balls are equally likely to be drawn

Problem 8.8
Repeat the previous problem but with each drawn ball to be put back into the urn.

Problem 8.9
An urn contains 16 black balls and 3 purple balls. Two balls are to be drawn one at a time without replacement. What is the probability of drawing out a black on the first draw and a purple on the second?

Problem 8.10
A board of trustees of a university consists of 8 men and 7 women. A committee of 3 must be selected at random and without replacement. The role of the committee is to select a new president for the university. Calculate the probability that the number of men selected exceeds the number of women selected.

Problem 8.11 ‡
A store has 80 modems in its inventory, 30 coming from Source A and the remainder from Source B. Of the modems from Source A, 20% are defective. Of the modems from Source B, 8% are defective.
Calculate the probability that exactly two out of a random sample of five modems from the store's inventory are defective.

Problem 8.12 ‡
From 27 pieces of luggage, an airline luggage handler damages a random sample of four. The probability that exactly one of the damaged pieces of luggage is insured is twice the probability that none of the damaged pieces are insured. Calculate the probability that exactly two of the four damaged pieces are insured.

Conditional Probability and Independence

In this chapter we introduce the concept of conditional probability. So far, the notation $\Pr(A)$ stands for the probability of A regardless of the occurrence of any other events. If the occurrence of an event B influences the probability of A then this new probability is called conditional probability.

9 Conditional Probabilities

We desire to know the probability of an event A conditional on the knowledge that another event B has occurred. The information the event B has occurred causes us to update the probabilities of other events in the sample space.
To illustrate, suppose you cast two dice; one red, and one green. Then the probability of getting two ones is $1/36$. However, if, after casting the dice, you ascertain that the green die shows a one (but know nothing about the red die), then there is a $1/6$ chance that both of them will be one. In other words, the probability of getting two ones changes if you have partial information, and we refer to this (altered) probability as **conditional probability**.
If the occurrence of the event A depends on the occurrence of B then the conditional probability will be denoted by $\Pr(A|B)$, read as the *probability of A given B*. Conditioning restricts the sample space to those outcomes which are in the set being conditioned on (in this case B). In this case,

$$\Pr(A|B) = \frac{\text{number of outcomes corresponding to event A and B}}{\text{number of outcomes of B}}.$$

Thus,

$$\Pr(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{\frac{n(A\cap B)}{n(S)}}{\frac{n(B)}{n(S)}} = \frac{\Pr(A \cap B)}{\Pr(B)}$$

provided that $P(B) > 0$.

Example 9.1
Let M denote the event "student is male" and let H denote the event "student is hispanic". In a class of 100 students suppose 60 are hispanic, and suppose that 10 of the hispanic students are males. Find the probability that a randomly chosen hispanic student is a male, that is, find $\Pr(M|H)$.

Solution.
Since 10 out of 100 students are both hispanic and male, $\Pr(M \cap H) = \frac{10}{100} = 0.1$. Also, 60 out of the 100 students are hispanic, so $\Pr(H) = \frac{60}{100} = 0.6$. Hence, $\Pr(M|H) = \frac{0.1}{0.6} = \frac{1}{6}$ ■

Using the formula

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

we can write

$$\Pr(A \cap B) = \Pr(A|B)\Pr(B) = \Pr(B|A)\Pr(A). \tag{9.1}$$

Example 9.2
The probability of a applicant to be admitted to a certain college is 0.8. The probability for a student in the college to live on campus is 0.6. What is the probability that an applicant will be admitted to the college and will be assigned a dormitory housing?

Solution.
The probability of the applicant being admitted and receiving dormitory housing is defined by

$$\begin{aligned}\Pr(\text{Accepted and Housing}) &= \Pr(\text{Housing}|\text{Accepted})\Pr(\text{Accepted}) \\ &= (0.6)(0.8) = 0.48 \blacksquare\end{aligned}$$

Equation (9.1) can be generalized to any finite number of events.

Theorem 9.1
Consider n events $A_1, A_2, \cdots, A_n$. Then

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

Proof.
The proof is by induction on $n \geq 2$. By Equation (9.1, the relation holds for $n = 2$. Suppose that the relation is true for $2, 3, \cdots, n$. We wish to establish

$$\Pr(A_1 \cap A_2 \cap \cdots \cap A_{n+1}) = \Pr(A_1)\Pr(A_2|A_1)\Pr(A_3|A_1 \cap A_2)\cdots\Pr(A_{n+1}|A_1 \cap A_2 \cap \cdots \cap A_n)$$

We have

$$\begin{aligned}
\Pr(A_1 \cap A_2 \cap \cdots \cap A_{n+1}) =& \Pr((A_1 \cap A_2 \cap \cdots \cap A_n) \cap A_{n+1})\\
=& \Pr(A_{n+1}|A_1 \cap A_2 \cap \cdots \cap A_n)\Pr(A_1 \cap A_2 \cap \cdots \cap A_n)\\
=& \Pr(A_{n+1}|A_1 \cap A_2 \cap \cdots \cap A_n)\Pr(A_1)\Pr(A_2|A_1)\Pr(A_3|A_1 \cap A_2)\\
& \cdots \Pr(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})\\
=& \Pr(A_1)\Pr(A_2|A_1)\Pr(A_3|A_1 \cap A_2)\cdots\Pr(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})\\
& \times \Pr(A_{n+1}|A_1 \cap A_2 \cap \cdots \cap A_n) \blacksquare
\end{aligned}$$

Example 9.3
Suppose 5 cards are drawn from a deck of 52 playing cards. What is the probability that all cards are the same suit, i.e. a flush?

Solution.
We must find

$$\Pr(\text{a flush}) = \Pr(\text{5 spades}) + \Pr(\text{ 5 hearts }) + \Pr(\text{ 5 diamonds }) + \Pr(\text{ 5 clubs })$$

Now, the probability of getting 5 spades is found as follows:

$$\begin{aligned}
\Pr(\text{5 spades}) &= \Pr(\text{1st card is a spade})\Pr(\text{2nd card is a spade}|\text{1st card is a spade})\\
&\times \cdots \times \Pr(\text{5th card is a spade}|\text{1st,2nd,3rd,4th cards are spades})\\
&= \tfrac{13}{52} \times \tfrac{12}{51} \times \tfrac{11}{50} \times \tfrac{10}{49} \times \tfrac{9}{48}
\end{aligned}$$

Since the above calculation is the same for any of the four suits,

$$\Pr(\text{a flush}) = 4 \times \tfrac{13}{52} \times \tfrac{12}{51} \times \tfrac{11}{50} \times \tfrac{10}{49} \times \tfrac{9}{48} \blacksquare$$

We end this section by showing that $\Pr(\cdot|A)$ satisfies the properties of ordinary probabilities.

Theorem 9.2
The function $B \to \Pr(B|A)$ defines a probability measure.

Proof.
1. Since $0 \leq \Pr(A \cap B) \leq \Pr(A)$, $0 \leq \Pr(B|A) \leq 1$.
2. $\Pr(S|A) = \frac{\Pr(S\cap A)}{\Pr(A)} = \frac{\Pr(A)}{\Pr(A)} = 1$.
3. Suppose that $B_1, B_2, \cdots$, are mutually exclusive events. Then $B_1 \cap A, B_2 \cap A, \cdots$, are mutually exclusive. Thus,

$$\begin{aligned}
\Pr(\bigcup_{n=1}^{\infty} B_n|A) =& \frac{\Pr\left(\bigcup_{n=1}^{\infty}(B_n \cap A)\right)}{\Pr(A)} \\
=& \frac{\sum_{n=1}^{\infty} \Pr(B_n \cap A)}{\Pr(A)} = \sum_{n=1}^{\infty} \Pr(B_n|A) \blacksquare
\end{aligned}$$

Thus, every theorem we have proved for an ordinary probability function holds for a conditional probability function. For example, we have

$$\Pr(B^c|A) = 1 - \Pr(B|A).$$

Prior and Posterior Probabilities
The probability $\Pr(A)$ is the probability of the event A prior to introducing new events that might affect A. It is known as the **prior probability** of A. When the occurrence of an event B will affect the event A then $\Pr(A|B)$ is known as the **posterior probability** of A.

Practice Problems

Problem 9.1 ‡
A public health researcher examines the medical records of a group of 937 men who died in 1999 and discovers that 210 of the men died from causes related to heart disease. Moreover, 312 of the 937 men had at least one parent who suffered from heart disease, and, of these 312 men, 102 died from causes related to heart disease.
Determine the probability that a man randomly selected from this group died of causes related to heart disease, given that neither of his parents suffered from heart disease.

Problem 9.2 ‡
An insurance company examines its pool of auto insurance customers and gathers the following information:

(i)	All customers insure at least one car.
(ii)	70% of the customers insure more than one car.
(iii)	20% of the customers insure a sports car.
(iv)	Of those customers who insure more than one car, 15% insure a sports car.

Calculate the probability that a randomly selected customer insures exactly one car and that car is not a sports car.

Problem 9.3 ‡
An actuary is studying the prevalence of three health risk factors, denoted by A, B, and C, within a population of women. For each of the three factors, the probability is 0.1 that a woman in the population has only this risk factor (and no others). For any two of the three factors, the probability is 0.12 that she has exactly these two risk factors (but not the other). The probability that a woman has all three risk factors, given that she has A and B, is $\frac{1}{3}$.
What is the probability that a woman has none of the three risk factors, given that she does not have risk factor A?

Problem 9.4
You are given $\Pr(A) = \frac{2}{5},\ \Pr(A \cup B) = \frac{3}{5},\ \Pr(B|A) = \frac{1}{4},\ \Pr(C|B) = \frac{1}{3}$, and $\Pr(C|A \cap B) = \frac{1}{2}$. Find $\Pr(A|B \cap C)$.

Problem 9.5
A pollster surveyed 100 people about watching the TV show "The big bang theory". The results of the poll are shown in the table.

	Yes	No	Total
Male	19	41	60
Female	12	28	40
Total	31	69	100

(a) What is the probability of a randomly selected individual is a male and watching the show?
(b) What is the probability of a randomly selected individual is a male?
(c) What is the probability of a randomly selected individual watches the show?
(d) What is the probability of a randomly selected individual watches the show, given that the individual is a male?
(e) What is the probability that a randomly selected individual watching the show is a male?

Problem 9.6
An urn contains 22 marbles: 10 red, 5 green, and 7 orange. You pick two at random without replacement. What is the probability that the first is red and the second is orange?

Problem 9.7
You roll two fair dice. Find the (conditional) probability that the sum of the two faces is 6 given that the two dice are showing different faces.

Problem 9.8
A machine produces small cans that are used for baked beans. The probability that the can is in perfect shape is 0.9. The probability of the can having an unnoticeable dent is 0.02. The probability that the can is obviously dented is 0.08. Produced cans get passed through an automatic inspection machine, which is able to detect obviously dented cans and discard them. What is the probability that a can that gets shipped for use will be of perfect shape?

Problem 9.9
An urn contains 225 white marbles and 15 black marbles. If we randomly pick (without replacement) two marbles in succession from the urn, what is the probability that they will both be black?

Problem 9.10
Find the probabilities of randomly drawing two kings in succession from an ordinary deck of 52 playing cards if we sample
(a) without replacement
(b) with replacement

Problem 9.11
A box of television tubes contains 20 tubes, of which five are defective. If three of the tubes are selected at random and removed from the box in succession without replacement, what is the probability that all three tubes are defective?

Problem 9.12
A study of texting and driving has found that 40% of all fatal auto accidents are attributed to texting drivers, 1% of all auto accidents are fatal, and drivers who text while driving are responsible for 20% of all accidents. Find the percentage of non-fatal accidents caused by drivers who do not text.

Problem 9.13
A TV manufacturer buys TV tubes from three sources. Source A supplies 50% of all tubes and has a 1% defective rate. Source B supplies 30% of all tubes and has a 2% defective rate. Source C supplies the remaining 20% of tubes and has a 5% defective rate.
(a) What is the probability that a randomly selected purchased tube is defective?
(b) Given that a purchased tube is defective, what is the probability it came from Source A? From Source B? From Source C?

Problem 9.14
In a certain town in the United States, 40% of the population are liberals and 60% are conservatives. The city council has proposed selling alcohol illegal in the town. It is known that 75% of conservatives and 30% of liberals support this measure.
(a) What is the probability that a randomly selected resident from the town will support the measure?
(b) If a randomly selected person does support the measure, what is the probability the person is a liberal?
(c) If a randomly selected person does not support the measure, what is the probability that he or she is a liberal?

10 Posterior Probabilities: Bayes' Formula

It is often the case that we know the probabilities of certain events conditional on other events, but what we would like to know is the "reverse". That is, given $\Pr(A|B)$ we would like to find $\Pr(B|A)$.
Bayes' formula is a simple mathematical formula used for calculating $\Pr(B|A)$ given $\Pr(A|B)$. We derive this formula as follows. Let A and B be two events. Then

$$A = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c).$$

Since the events $A \cap B$ and $A \cap B^c$ are mutually exclusive, we can write

$$\begin{aligned} \Pr(A) &= \Pr(A \cap B) + \Pr(A \cap B^c) \\ &= \Pr(A|B)\Pr(B) + \Pr(A|B^c)\Pr(B^c) \end{aligned} \tag{10.1}$$

Example 10.1
The completion of a highway construction may be delayed because of a projected storm. The probabilities are 0.60 that there will be a storm, 0.85 that the construction job will be completed on time if there is no storm, and 0.35 that the construction will be completed on time if there is a storm. What is the probability that the construction job will be completed on time?

Solution.
Let A be the event that the construction job will be completed on time and B is the event that there will be a storm. We are given $\Pr(B) = 0.60$, $\Pr(A|B^c) = 0.85$, and $\Pr(A|B) = 0.35$. From Equation (10.1) we find

$$\Pr(A) = \Pr(B)\Pr(A|B) + \Pr(B^c)\Pr(A|B^c) = (0.60)(0.35) + (0.4)(0.85) = 0.55 \blacksquare$$

From Equation (10.1) we can get Bayes' formula:

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A \cap B) + \Pr(A \cap B^c)} = \frac{\Pr(A|B)\Pr(B)}{\Pr(A|B)\Pr(B) + \Pr(A|B^c)\Pr(B^c)} \tag{10.2}$$

Example 10.2
A small manufacturing company uses two machines A and B to make shirts. Observation shows that machine A produces 10% of the total production of shirts while machine B produces 90% of the total production of shirts. Assuming that 1% of all the shirts produced by A are defective while 5% of

all the shirts produced by B are defective, find the probability that a shirt taken at random from a day's production was made by machine A, given that it is defective.

Solution.
We are given $\Pr(A) = 0.1,\ \Pr(B) = 0.9,\ \Pr(D|A) = 0.01$, and $\Pr(D|B) = 0.05$. We want to find $\Pr(A|D)$. Using Bayes' formula we find

$$\begin{aligned}\Pr(A|D) =&\frac{\Pr(A\cap D)}{\Pr(D)} = \frac{\Pr(D|A)\Pr(A)}{\Pr(D|A)\Pr(A)+\Pr(D|B)\Pr(B)}\\ =&\frac{(0.01)(0.1)}{(0.01)(0.1)+(0.05)(0.9)} \approx 0.0217 \blacksquare\end{aligned}$$

Example 10.3
A credit card company offers two types of cards: a basic card (B) and a gold card (G). Over the past year, 40% of the cards issued have been of the basic type. Of those getting the basic card, 30% enrolled in an identity theft plan, whereas 50% of all gold cards holders do so. If you learn that a randomly selected cardholder has an identity theft plan, how likely is it that he/she has a basic card?

Solution.
Let I denote the identity theft plan. We are given $\Pr(B) = 0.4, \Pr(G) = 0.6,\ \Pr(I|B) = 0.3$, and $\Pr(I|G) = 0.5$. By Bayes' formula we have

$$\begin{aligned}\Pr(B|I) =&\frac{\Pr(B\cap I)}{\Pr(I)} = \frac{\Pr(I|B)\Pr(B)}{\Pr(I|B)\Pr(B)+\Pr(I|G)\Pr(G)}\\ =&\frac{(0.3)(0.4)}{(0.3)(0.4)+(0.5)(0.6)} = 0.286 \blacksquare\end{aligned}$$

Formula (10.2) is a special case of the more general result:

Theorem 10.1 *(Bayes' formula)*
Suppose that the sample space S is the union of mutually exclusive events $H_1, H_2, \cdots, H_n$ with $\Pr(H_i) > 0$ for each i. Then for any event A and $1 \leq i \leq n$ we have

$$\Pr(H_i|A) = \frac{\Pr(A|H_i)\Pr(H_i)}{\Pr(A)}$$

where

$$\Pr(A) = \Pr(H_1)\Pr(A|H_1) + \Pr(H_2)\Pr(A|H_2) + \cdots + \Pr(H_n)\Pr(A|H_n).$$

Proof.
First note that

$$\begin{aligned}\Pr(A) =&\Pr(A\cap S) = \Pr(A\cap(\bigcup_{i=1}^n H_i))\\ =&\Pr(\bigcup_{i=1}^n (A\cap H_i))\\ =&\sum_{i=1}^n \Pr(A\cap H_i) = \sum_{i=1}^n \Pr(A|H_i)\Pr(H_i)\end{aligned}$$

Hence,

$$\Pr(H_i|A) = \frac{\Pr(A|H_i)\Pr(H_i)}{\Pr(A)} = \frac{\Pr(A|H_i)\Pr(H_i)}{\sum_{i=1}^n \Pr(A|H_i)\Pr(H_i)} \blacksquare$$

Example 10.4
A survey about a measure to legalize medical marijuanah is taken in three states: Kentucky, Maine and Arkansas. In Kentucky, 50% of voters support the measure, in Maine, 60% of the voters support the measure, and in Arkansas, 35% of the voters support the measure. Of the total population of the three states, 40% live in Kentucky , 25% live in Maine, and 35% live in Arkansas. Given that a voter supports the measure, what is the probability that he/she lives in Maine?

Solution.
Let L_I denote the event that a voter lives in state I, where I = K (Kentucky), M (Maine), or A (Arkansas). Let S denote the event that a voter supports the measure. We want to find $\Pr(L_M|S)$. By Bayes' formula we have

$$\begin{aligned}\Pr(L_M|S) =& \frac{\Pr(S|L_M)\Pr(L_M)}{\Pr(S|L_K)\Pr(L_K)+\Pr(S|L_M)\Pr(L_M)+\Pr(S|L_A)\Pr(L_A)}\\ =& \frac{(0.6)(0.25)}{(0.5)(0.4)+(0.6)(0.25)+(0.35)(0.35)} \approx 0.3175 \blacksquare\end{aligned}$$

Example 10.5
Passengers in Little Rock Airport rent cars from three rental companies: 60% from Avis, 30% from Enterprise, and 10% from National. Past statistics show that 9% of the cars from Avis, 20% of the cars from Enterprise , and 6% of the cars from National need oil change. If a rental car delivered to a passenger needs an oil change, what is the probability that it came from Enterprise?

Solution.
Define the events

$$\begin{aligned}A =& \text{car comes from Avis}\\ E =& \text{car comes from Enterprise}\\ N =& \text{car comes from National}\\ O =& \text{car needs oil change}\end{aligned}$$

Then

$$\begin{array}{ccc}\Pr(A) = 0.6 & \Pr(E) = 0.3 & \Pr(N) = 0.1\\ \Pr(O|A) = 0.09 & \Pr(O|E) = 0.2 & \Pr(O|N) = 0.06\end{array}$$

From Bayes' theorem we have

$$\begin{aligned}\Pr(E|O) =& \frac{\Pr(O|E)\Pr(E)}{\Pr(O|A)\Pr(A) + \Pr(O|E)\Pr(E) + \Pr(O|N)\Pr(N)}\\ =& \frac{0.2 \times 0.3}{0.09 \times 0.6 + 0.2 \times 0.3 + 0.06 \times 0.1} = 0.5 \blacksquare\end{aligned}$$

Example 10.6
A toy factory produces its toys with three machines A, B, and C. From the total production, 50% are produced by machine A, 30% by machine B, and 20% by machine C. Past statistics show that 4% of the toys produced by machine A are defective, 2% produced by machine B are defective, and 4% of the toys produced by machine C are defective.
(a) What is the probability that a randomly selected toy is defective?
(b) If a randomly selected toy was found to be defective, what is the probability that this toy was produced by machine A?

Solution.
Let D be the event that the selected product is defective. Then, $\Pr(A) = 0.5, \Pr(B) = 0.3, \Pr(C) = 0.2, \Pr(D|A) = 0.04,\ \Pr(D|B) = 0.02, \Pr(D|C) = 0.04$. We have

$$\begin{aligned}\Pr(D) =& \Pr(D|A)\Pr(A) + \Pr(D|B)\Pr(B) + \Pr(D|C)\Pr(C)\\ =& (0.04)(0.50) + (0.02)(0.30) + (0.04)(0.20) = 0.034\end{aligned}$$

(b) By Bayes' theorem, we find

$$\Pr(A|D) = \frac{\Pr(D|A)\Pr(A)}{\Pr(D)} = \frac{(0.04)(0.50)}{0.034} \approx 0.5882 \blacksquare$$

Example 10.7
A group of traffic violators consists of 45 men and 15 women. The men have probability 1/2 for being ticketed for crossing a red light while the women have probability 1/3 for the same offense.
(a) Suppose you choose at random a person from the group. What is the probability that the person will be ticketed for crossing a red light?
(b) Determine the conditional probability that you chose a woman given that the person you chose was being ticketed for crossing the red light.

Solution.
Let

$$\begin{aligned} W =&\{\text{the one selected is a woman}\}\\ M =&\{\text{the one selected is a man}\}\\ T =&\{\text{the one selected is ticketed for crossing a red light}\}\end{aligned}$$

(a) We are given the following information: $\Pr(W) = \frac{15}{60} = \frac{1}{4}, \Pr(M) = \frac{3}{4},\ \Pr(T|W) = \frac{1}{3}$, and $\Pr(T|M) = \frac{1}{2}$. We have,

$$\Pr(T) = \Pr(T|W)\Pr(W) + \Pr(T|M)\Pr(M) = \frac{11}{24}.$$

(b) Using Bayes' theorem we find

$$\Pr(W|T) = \frac{\Pr(T|W)\Pr(W)}{\Pr(T)} = \frac{(1/3)(1/4)}{(11/24)} = \frac{2}{11} \blacksquare$$

Practice Problems

Problem 10.1
An insurance company believes that auto drivers can be divided into two categories: those who are a high risk for accidents and those who are low risk. Past statistics show that the probability for a high risk driver to have an accident within a one-year period is 0.4, whereas this probability is 0.2 for a low risk driver.
(a) If we assume that 30% of the population is high risk, what is the probability that a new policyholder will have an accident within a year of purchasing a policy?
(b) Suppose that a new policyholder has an accident within a year of purchasing a policy. What is the probability that he or she is high risk?

Problem 10.2 ‡
An auto insurance company insures drivers of all ages. An actuary compiled the following statistics on the company's insured drivers:

Age of Driver	Probability of Accident	Portion of Company's Insured Drivers
16 - 20	0.06	0.08
21 - 30	0.03	0.15
31 - 65	0.02	0.49
66 - 99	0.04	0.28

A randomly selected driver that the company insures has an accident. Calculate the probability that the driver was age 16-20.

Problem 10.3 ‡
An insurance company issues life insurance policies in three separate categories: standard, preferred, and ultra-preferred. Of the company's policyholders, 50% are standard, 40% are preferred, and 10% are ultra-preferred. Each standard policyholder has probability 0.010 of dying in the next year, each preferred policyholder has probability 0.005 of dying in the next year, and each ultra-preferred policyholder has probability 0.001 of dying in the next year.
A policyholder dies in the next year. What is the probability that the deceased policyholder was ultra-preferred?

Problem 10.4 ‡
Upon arrival at a hospital's emergency room, patients are categorized according to their condition as critical, serious, or stable. In the past year:

(i) 10% of the emergency room patients were critical;
(ii) 30% of the emergency room patients were serious;
(iii) the rest of the emergency room patients were stable;
(iv) 40% of the critical patients died;
(v) 10% of the serious patients died; and
(vi) 1% of the stable patients died.

Given that a patient survived, what is the probability that the patient was categorized as serious upon arrival?

Problem 10.5 ‡
A health study tracked a group of persons for five years. At the beginning of the study, 20% were classified as heavy smokers, 30% as light smokers, and 50% as nonsmokers.
Results of the study showed that light smokers were twice as likely as nonsmokers to die during the five-year study, but only half as likely as heavy smokers.
A randomly selected participant from the study died over the five-year period. Calculate the probability that the participant was a heavy smoker.

Problem 10.6 ‡
An actuary studied the likelihood that different types of drivers would be involved in at least one collision during any one-year period. The results of the study are presented below.

Type of driver	Percentage of all drivers	Probability of at least one collision
Teen	8%	0.15
Young adult	16%	0.08
Midlife	45%	0.04
Senior	31%	0.05
Total	100%	

Given that a driver has been involved in at least one collision in the past year, what is the probability that the driver is a young adult driver?

Problem 10.7 ‡
A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence of the disease 0.5% of the time when the disease is not present. One percent of the population actually has the disease.
Calculate the probability that a person has the disease given that the test indicates the presence of the disease.

Problem 10.8 ‡
The probability that a randomly chosen male has a circulation problem is 0.25 . Males who have a circulation problem are twice as likely to be smokers as those who do not have a circulation problem.
What is the conditional probability that a male has a circulation problem, given that he is a smoker?

Problem 10.9 ‡
A study of automobile accidents produced the following data:

Model year	Proportion of all vehicles	Probability of involvement in an accident
1997	0.16	0.05
1998	0.18	0.02
1999	0.20	0.03
Other	0.46	0.04

An automobile from one of the model years 1997, 1998, and 1999 was involved in an accident. Determine the probability that the model year of this automobile is 1997.

Problem 10.10
A study was conducted about the excessive amounts of pollutants emitted by cars in a certain town. The study found that 25% of all cars emit excessive amounts of pollutants. The probability for a car emiting excessive amounts of pollutants to fail the town's vehicular emission test is found to be 0.99. Cars who do not emit excessive amounts of pollutants have a probability of 0.17 to fail to emission test. A car is selected at random. What is the probability that the car emits excessive amounts of pollutants given that it failed the emission test?

Problem 10.11
A medical agency is conducting a study about injuries resulted from activities for a group of people. For this group, 50% were skiing, 30% were hiking, and 20% were playing soccer. The (conditional) probability of a person getting injured from skiing is 30% , it is 10% from hiking, and 20% from playing soccer.
(a) What is the probability for a randomly selected person in the group for getting injured?
(b) Given that a person is injured, what is the probability that his injuries are due to skiing?

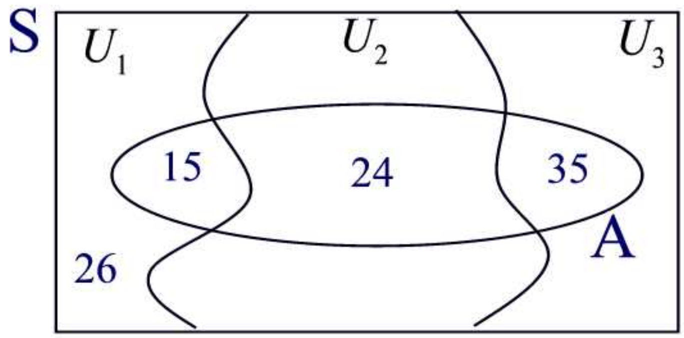
Problem 10.12
A written driving test is graded either pass or fail. A randomly chosen person from a driving class has a 40% chance of knowing the material well. If the person knows the material well, the probability for this person to pass the written test is 0.8. For a person not knowing the material well, the probability is 0.4 for passing the test.
(a) What is the probability of a randomly chosen person from the class for passing the test?
(b) Given that a person in the class passes the test, what is the probability that this person knows the material well?

Problem 10.13 ‡
Ten percent of a company's life insurance policyholders are smokers. The rest are nonsmokers. For each nonsmoker, the probability of dying during the year is 0.01. For each smoker, the probability of dying during the year is 0.05.
Given that a policyholder has died, what is the probability that the policyholder was a smoker?

Problem 10.14
A prerequisite for students to take a probability class is to pass calculus. A study of correlation of grades for students taking calculus and probability was conducted. The study shows that 25% of all calculus students get an A, and that students who had an A in calculus are 50% more likely to get an A in probability as those who had a lower grade in calculus. If a student who received an A in probability is chosen at random, what is the probability that he/she also received an A in calculus?

Problem 10.15
A group of people consists of 70 men and 70 women. Seven men and ten women are found to be color-blind.
(a) What is the probability that a randomly selected person is color-blind?
(b) If the randomly selected person is color-blind, what is the probability that the person is a man?

Problem 10.16
Calculate $\Pr(U_1|A)$.



Problem 10.17
The probability that a person with certain symptoms has prostate cancer is 0.8. A PSA test used to confirm this diagnosis gives positive results for 90% of those who have the disease, and 5% of those who do not have the disease. What is the probability that a person who reacts positively to the test actually has the disease ?

11 Independent Events

Intuitively, when the occurrence of an event B has no influence on the probability of occurrence of an event A then we say that the two events are independent. For example, in the experiment of tossing two coins, the first toss has no effect on the second toss. In terms of conditional probability, two events A and B are said to be **independent** if and only if

$$\Pr(A|B) = \Pr(A).$$

We next introduce the two most basic theorems regarding independence.

Theorem 11.1
A and B are independent events if and only if $\Pr(A \cap B) = \Pr(A)\Pr(B)$.

Proof.
A and B are independent if and only if $\Pr(A|B) = \Pr(A)$ and this is equivalent to

$$\Pr(A \cap B) = \Pr(A|B)\Pr(B) = \Pr(A)\Pr(B) \blacksquare$$

Example 11.1
Show that $\Pr(A|B) > \Pr(A)$ if and only if $\Pr(A^c|B) < \Pr(A^c)$. We assume that $0 < \Pr(A) < 1$ and $0 < \Pr(B) < 1$

Solution.
We have

$$\begin{aligned}
\Pr(A|B) > \Pr(A) \Leftrightarrow & \frac{\Pr(A \cap B)}{\Pr(B)} > \Pr(A)\\
\Leftrightarrow & \Pr(A \cap B) > \Pr(A)\Pr(B)\\
\Leftrightarrow & \Pr(B) - \Pr(A \cap B) < \Pr(B) - \Pr(A)\Pr(B)\\
\Leftrightarrow & \Pr(A^c \cap B) < \Pr(B)(1 - \Pr(A))\\
\Leftrightarrow & \Pr(A^c \cap B) < \Pr(B)\Pr(A^c)\\
\Leftrightarrow & \frac{\Pr(A^c \cap B)}{\Pr(B)} < \Pr(A^c)\\
\Leftrightarrow & \Pr(A^c|B) < \Pr(A^c) \blacksquare
\end{aligned}$$

Example 11.2
A coal exploration company is set to look for coal mines in two states Virginia and New Mexico. Let A be the event that a coal mine is found in Virginia and B the event that a coal mine is found in New Mexico. Suppose that A and B are independent events with $\Pr(A) = 0.4$ and $\Pr(B) = 0.7$. What is the probability that at least one coal mine is found in one of the states?

Solution.
The probability that at least one coal mine is found in one of the two states is $\Pr(A \cup B)$. Thus,

$$\begin{aligned}\Pr(A \cup B) =&\Pr(A) + \Pr(B) - \Pr(A \cap B)\\ =&\Pr(A) + \Pr(B) - \Pr(A)\Pr(B)\\ =&0.4 + 0.7 - 0.4(0.7) = 0.82 \blacksquare\end{aligned}$$

Example 11.3
Let A and B be two independent events such that $\Pr(B|A \cup B) = \frac{2}{3}$ and $\Pr(A|B) = \frac{1}{2}$. What is $\Pr(B)$?

Solution.
First, note that by indepedence we have

$$\frac{1}{2} = \Pr(A|B) = \Pr(A).$$

Next,

$$\begin{aligned}\Pr(B|A \cup B) =&\frac{\Pr(B)}{\Pr(A \cup B)}\\ =&\frac{\Pr(B)}{\Pr(A) + \Pr(B) - \Pr(A \cap B)}\\ =&\frac{\Pr(B)}{\Pr(A) + \Pr(B) - \Pr(A)\Pr(B)}.\end{aligned}$$

Thus,

$$\frac{2}{3} = \frac{\Pr(B)}{\frac{1}{2} + \frac{\Pr(B)}{2}}$$

Solving this equation for $\Pr(B)$ we find $\Pr(B) = \frac{1}{2}$ $\blacksquare$

Theorem 11.2
If A and B are independent then so are A and B^c.

Proof.
First note that A can be written as the union of two mutually exclusive events: $A = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c)$. Thus, $\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B^c)$. It follows that

$$\begin{aligned}\Pr(A \cap B^c) =& \Pr(A) - \Pr(A \cap B)\\ =& \Pr(A) - \Pr(A)\Pr(B)\\ =& \Pr(A)(1 - \Pr(B)) = \Pr(A)\Pr(B^c) \blacksquare\end{aligned}$$

Example 11.4
Show that if A and B are independent so are A^c and B^c.

Solution.
Using De Morgan's formula we have

$$\begin{aligned}\Pr(A^c \cap B^c) =& 1 - \Pr(A \cup B) = 1 - [\Pr(A) + \Pr(B) - \Pr(A \cap B)]\\ =& [1 - \Pr(A)] - \Pr(B) + \Pr(A)\Pr(B)\\ =& \Pr(A^c) - \Pr(B)[1 - \Pr(A)] = \Pr(A^c) - \Pr(B)\Pr(A^c)\\ =& \Pr(A^c)[1 - \Pr(B)] = \Pr(A^c)\Pr(B^c) \blacksquare\end{aligned}$$

When the outcome of one event affects the outcome of a second event, the events are said to be **dependent.** The following is an example of events that are not independent.

Example 11.5
Draw two cards from a deck. Let A = "The first card is a spade," and B = "The second card is a spade." Show that A and B are dependent.

Solution.
We have $\Pr(A) = \Pr(B) = \frac{13}{52} = \frac{1}{4}$ and

$$\Pr(A \cap B) = \frac{13 \cdot 12}{52 \cdot 51} < \left(\frac{1}{4}\right)^2 = \Pr(A)\Pr(B).$$

By Theorem 11.1 the events A and B are dependent ■

The definition of independence for a finite number of events is defined as follows: Events $A_1, A_2, \cdots, A_n$ are said to be **mutually independent** or simply **independent** if for any $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ we have

$$\Pr(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \Pr(A_{i_1})\Pr(A_{i_2}) \cdots \Pr(A_{i_k})$$

In particular, three events A, B, C are independent if and only if

$$\begin{aligned}
\Pr(A \cap B) =& \Pr(A)\Pr(B)\\
\Pr(A \cap C) =& \Pr(A)\Pr(C)\\
\Pr(B \cap C) =& \Pr(B)\Pr(C)\\
\Pr(A \cap B \cap C) =& \Pr(A)\Pr(B)\Pr(C)
\end{aligned}$$

Example 11.6
Consider the experiment of tossing a coin n times. Let $A_i =$ "the i$^{\text{th}}$ coin shows Heads". Show that $A_1, A_2, \cdots, A_n$ are independent.

Solution.
For any $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ we have $\Pr(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \frac{1}{2^k}$. But $\Pr(A_i) = \frac{1}{2}$. Thus, $\Pr(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \Pr(A_{i_1})\Pr(A_{i_2}) \cdots \Pr(A_{i_k})$ ■

Example 11.7
In a clinic laboratory, the probability that a blood sample shows cancerous cells is 0.05. Four blood samples are tested, and the samples are independent.
(a) What is the probability that none shows cancerous cells?
(b) What is the probability that exactly one sample shows cancerous cells?
(c) What is the probability that at least one sample shows cancerous cells?

Solution.
Let H_i denote the event that the i^{th} sample contains cancerous cells for $i = 1, 2, 3, 4$.
The event that none contains cancerous cells is equivalent to $H_1^c \cap H_2^c \cap H_3^c \cap H_4^c$. So, by independence, the desired probability is

$$\begin{aligned}
\Pr(H_1^c \cap H_2^c \cap H_3^c \cap H_4^c) =& \Pr(H_1^c)\Pr(H_2^c)\Pr(H_3^c)\Pr(H_4^c)\\
=&(1 - 0.05)^4 = 0.8145
\end{aligned}$$

(b) Let

$$\begin{aligned}A_1 =& H_1 \cap H_2^c \cap H_3^c \cap H_4^c \\ A_2 =& H_1^c \cap H_2 \cap H_3^c \cap H_4^c \\ A_3 =& H_1^c \cap H_2^c \cap H_3 \cap H_4^c \\ A_4 =& H_1^c \cap H_2^c \cap H_3^c \cap H_4\end{aligned}$$

Then, the requested probability is the probability of the union $A_1 \cup A_2 \cup A_3 \cup A_4$ and these events are mutually exclusive. Also, by independence, $\Pr(A_i) = (0.95)^3(0.05) = 0.0429, i = 1, 2, 3, 4$. Therefore, the answer is $4(0.0429) = 0.1716$.

(c) Let B be the event that no sample contains cancerous cells. The event that at least one sample contains cancerous cells is the complement of B, i.e. B^c. By part (a), it is known that $\Pr(B) = 0.8145$. So, the requested probability is

$$\Pr(B^c) = 1 - \Pr(B) = 1 - 0.8145 = 0.1855 \blacksquare$$

Example 11.8
Find the probability of getting four sixes and then another number in five random rolls of a balanced die.

Solution.
Because the events are independent, the probability in question is

$$\frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{7776} \blacksquare$$

A collection of events $A_1, A_2, \cdots, A_n$ are said to be **pairwise independent** if and only if $\Pr(A_i \cap A_j) = \Pr(A_i)\Pr(A_j)$ for any $i \neq j$ where $1 \leq i, j \leq n$. Pairwise independence does not imply mutual independence as the following example shows.

Example 11.9
Consider the experiment of flipping two fair coins. Consider the three events: A = the first coin shows heads; B = the second coin shows heads, and C = the two coins show the same result. Show that these events are pairwise independent, but not independent.

Solution.
Note that $A = \{H,H),(H,T)\},\ B = \{(H,H),(T,H)\},\ C = \{(H,H),(T,T)\}$.
We have

$$\begin{aligned}\Pr(A\cap B) =&\Pr(\{(H,H)\}) = \frac{1}{4} = \frac{2}{4}\cdot\frac{2}{4} = \Pr(A)\Pr(B)\\ \Pr(A\cap C) =&\Pr(\{(H,H)\}) = \frac{1}{4} = \frac{2}{4}\cdot\frac{2}{4} = \Pr(A)\Pr(C)\\ \Pr(B\cap C) =&\Pr(\{(H,H)\}) = \frac{1}{4} = \frac{2}{4}\cdot\frac{2}{4} = \Pr(B)\Pr(C)\end{aligned}$$

Hence, the events A, B, and C are pairwise independent. On the other hand

$$\Pr(A\cap B\cap C) = \Pr(\{(H,H)\}) = \frac{1}{4} \neq \frac{2}{4}\cdot\frac{2}{4}\cdot\frac{2}{4} = \Pr(A)\Pr(B)\Pr(C)$$

so that A, B, and C are not independent ■

Practice Problems

Problem 11.1
Determine whether the events are independent or dependent.
(a) Selecting a marble from an urn and then choosing a second marble from the same urn without replacing the first marble.
(b) Rolling a die and spinning a spinner.

Problem 11.2
Amin and Nadia are allowed to have one topping on their ice cream. The choices of toppings are Butterfingers, M and M, chocolate chips, Gummy Bears, Kit Kat, Peanut Butter, and chocolate syrup. If they choose at random, what is the probability that they both choose Kit Kat as a topping?

Problem 11.3
You randomly select two cards from a standard 52-card deck. What is the probability that the first card is not a face card (a king, queen, jack, or an ace) and the second card is a face card if
(a) you replace the first card before selecting the second, and
(b) you do not replace the first card?

Problem 11.4
Marlon, John, and Steve are given the choice for only one topping on their personal size pizza. There are 10 toppings to choose from. What is the probability that each of them orders a different topping?

Problem 11.5 ‡
One urn contains 4 red balls and 6 blue balls. A second urn contains 16 red balls and x blue balls. A single ball is drawn from each urn. The probability that both balls are the same color is 0.44 .
Calculate x.

Problem 11.6 ‡
An actuary studying the insurance preferences of automobile owners makes the following conclusions:
(i) An automobile owner is twice as likely to purchase a collision coverage as opposed to a disability coverage.
(ii) The event that an automobile owner purchases a collision coverage is independent of the event that he or she purchases a disability coverage.

(iii) The probability that an automobile owner purchases both collision and disability coverages is 0.15.
What is the probability that an automobile owner purchases neither collision nor disability coverage?

Problem 11.7 ‡
An insurance company pays hospital claims. The number of claims that include emergency room or operating room charges is 85% of the total number of claims. The number of claims that do not include emergency room charges is 25% of the total number of claims. The occurrence of emergency room charges is independent of the occurrence of operating room charges on hospital claims.
Calculate the probability that a claim submitted to the insurance company includes operating room charges.

Problem 11.8
Let $S = \{1, 2, 3, 4\}$ with each outcome having equal probability $\frac{1}{4}$ and define the events $A = \{1, 2\},\ B = \{1, 3\}$, and $C = \{1, 4\}$. Show that the three events are pairwise independent but not independent.

Problem 11.9
Assume A and B are independent events with $\Pr(A) = 0.2$ and $\Pr(B) = 0.3$. Let C be the event that neither A nor B occurs, let D be the event that exactly one of A or B occurs. Find $\Pr(C)$ and $\Pr(D)$.

Problem 11.10
Suppose A, B, and C are mutually independent events with probabilities $\Pr(A) = 0.5,\ \Pr(B) = 0.8$, and $\Pr(C) = 0.3$. Find the probability that at least one of these events occurs.

Problem 11.11
Suppose A, B, and C are mutually independent events with probabilities $\Pr(A) = 0.5, \Pr(B) = 0.8$, and $\Pr(C) = 0.3$. Find the probability that exactly two of the events A, B, C occur.

Problem 11.12
If events A, B, and C are independent, show that
(a) A and $B \cap C$ are independent
(b) A and $B \cup C$ are independent

Problem 11.13
Suppose you flip a nickel, a dime and a quarter. Each coin is fair, and the flips of the different coins are independent. Let A be the event "the total value of the coins that came up heads is at least 15 cents". Let B be the event "the quarter came up heads". Let C be the event "the total value of the coins that came up heads is divisible by 10 cents".
(a) Write down the sample space, and list the events A, B, and C.
(b) Find $\Pr(A), \Pr(B)$ and $\Pr(C)$.
(c) Compute $\Pr(B|A)$.
(d) Are B and C independent? Explain.

Problem 11.14 ‡
Workplace accidents are categorized into three groups: minor, moderate and severe. The probability that a given accident is minor is 0.5, that it is moderate is 0.4, and that it is severe is 0.1. Two accidents occur independently in one month.
Calculate the probability that neither accident is severe and at most one is moderate.

Problem 11.15
Among undergraduate students living on a college campus, 20% have an automobile. Among undergraduate students living off campus, 60% have an automobile. Among undergraduate students, 30% live on campus. Give the probabilities of the following events when a student is selected at random:
(a) Student lives off campus
(b) Student lives on campus and has an automobile
(c) Student lives on campus and does not have an automobile
(d) Student lives on campus or has an automobile
(e) Student lives on campus given that he/she does not have an automobile.

12 Odds and Conditional Probability

What's the difference between probabilities and odds? To answer this question, let's consider a game that involves rolling a die. If one gets the face 1 then he wins the game, otherwise he loses. The probability of winning is $\frac{1}{6}$ whereas the probability of losing is $\frac{5}{6}$. The odds of winning is 1:5(read 1 to 5). This expression means that the probability of losing is five times the probability of winning. Thus, probabilities describe the frequency of a favorable result in relation to all possible outcomes whereas the **odds in favor** of an event compare the favorable outcomes to the unfavorable outcomes. More formally,

$$\text{odds in favor} = \frac{\text{favorable outcomes}}{\text{unfavorable outcomes}}$$

If E is the event of all favorable outcomes then its complementary, E^c, is the event of unfavorable outcomes. Hence,

$$\text{odds in favor} = \frac{n(E)}{n(E^c)}$$

Also, we define the **odds against** an event as

$$\text{odds against} = \frac{\text{unfavorable outcomes}}{\text{favorable outcomes}} = \frac{n(E^c)}{n(E)}$$

Any probability can be converted to odds, and any odds can be converted to a probability.

Converting Odds to Probability

Suppose that the odds in favor for an event E is a:b. Thus, $n(E) = ak$ and $n(E^c) = bk$ where k is a positive integer. Since $S = E \cup E^c$ and $E \cap E^c = \emptyset$, by Theorem 2.3(b) we have $n(S) = n(E) + n(E^c)$. Therefore,

$$\Pr(E) = \frac{n(E)}{n(S)} = \frac{n(E)}{n(E)+n(E^c)} = \frac{ak}{ak+bk} = \frac{a}{a+b}$$

and

$$\Pr(E^c) = \frac{n(E^c)}{n(S)} = \frac{n(E^c)}{n(E)+n(E^c)} = \frac{bk}{ak+bk} = \frac{b}{a+b}$$

Example 12.1

If the odds in favor of an event E is 5:4, compute $\Pr(E)$ and $\Pr(E^c)$.

Solution.

We have

$$\Pr(E) = \frac{5}{5+4} = \frac{5}{9} \quad \text{and} \quad \Pr(E^c) = \frac{4}{5+4} = \frac{4}{9} \blacksquare$$

Converting Probability to Odds
Given $\Pr(E)$, we want to find the odds in favor of E and the odds against E. The odds in favor of E are

$$\begin{aligned} \frac{n(E)}{n(E^c)} &= \frac{n(E)}{n(S)} \cdot \frac{n(S)}{n(E^c)} \\ &= \frac{\Pr(E)}{\Pr(E^c)} \\ &= \frac{\Pr(E)}{1 - \Pr(E)} \end{aligned}$$

and the odds against E are

$$\frac{n(E^c)}{n(E)} = \frac{1 - \Pr(E)}{\Pr(E)}$$

Example 12.2
For each of the following, find the odds in favor of the event's occurring:
(a) Rolling a number less than 5 on a die.
(b) Tossing heads on a fair coin.
(c) Drawing an ace from an ordinary 52-card deck.

Solution.
(a) The probability of rolling a number less than 5 is $\frac{4}{6}$ and that of rolling 5 or 6 is $\frac{2}{6}$. Thus, the odds in favor of rolling a number less than 5 is $\frac{4}{6} \div \frac{2}{6} = \frac{2}{1}$ or 2:1
(b) Since $\Pr(H) = \frac{1}{2}$ and $\Pr(T) = \frac{1}{2}$, the odds in favor of getting heads is $\left(\frac{1}{2}\right) \div \left(\frac{1}{2}\right)$ or 1:1
(c) We have $\Pr(\text{ace}) = \frac{4}{52}$ and $\Pr(\text{not an ace}) = \frac{48}{52}$ so that the odds in favor of drawing an ace is $\left(\frac{4}{52}\right) \div \left(\frac{48}{52}\right) = \frac{1}{12}$ or 1:12 $\blacksquare$

Remark 12.1
A probability such as $\Pr(E) = \frac{5}{6}$ is just a ratio. The exact number of favorable outcomes and the exact total of all outcomes are not necessarily known.

Practice Problems

Problem 12.1
If the probability of a boy being born is $\frac{1}{2}$, and a family plans to have four children, what are the odds against having all boys?

Problem 12.2
If the odds against Nadia's winning first prize in a chess tournament are 3:5, what is the probability that she will win first prize?

Problem 12.3
What are the odds in favor of getting at least two heads if a fair coin is tossed three times?

Problem 12.4
If the probability of snow for the day is 60%, what are the odds against snowing?

Problem 12.5
On a tote board at a race track, the odds for Smarty Harper are listed as 26:1. Tote boards list the odds that the horse will lose the race. If this is the case, what is the probability of Smarty Harper's winning the race?

Problem 12.6
If a die is tossed, what are the odds in favor of the following events?
(a) Getting a 4
(b) Getting a prime
(c) Getting a number greater than 0
(d) Getting a number greater than 6.

Problem 12.7
Find the odds against E if $\Pr(E) = \frac{3}{4}$.

Problem 12.8
Find $\Pr(E)$ in each case.
(a) The odds in favor of E are 3:4
(b) The odds against E are 7:3

Discrete Random Variables

This chapter is one of two chapters dealing with random variables. After introducing the notion of a random variable, we discuss discrete random variables. Continuous random variables are left to the next chapter.

13 Random Variables

By definition, a **random variable** X is a function with domain the sample space and range a subset of the real numbers. For example, in rolling two dice X might represent the sum of the points on the two dice. Similarly, in taking samples of college students X might represent the number of hours per week a student studies, a student's GPA, or a student's height.
The notation $X(s) = x$ means that x is the value associated with the outcome s by the random variable X.
There are three types of random variables: discrete random variables, continuous random variables, and mixed random variables.
A **discrete** is a random variable whose range is either finite or countably infinite. A **continuous** random variable is a random variable whose range is an interval in $\mathbb{R}$. A **mixed** random variable is partially discrete and partially continuous.
In this chapter we will just consider discrete random variables.

Example 13.1
State whether the random variables are discrete, continuous or mixed.
(a) A coin is tossed ten times. The random variable X is the number of tails that are noted.
(b) A light bulb is burned until it burns out. The random variable Y is its lifetime in hours.

(c) $Z : (0,1) \to \mathbb{R}$ where

$$Z(s) = \begin{cases} 1 - s, & 0 < s < \frac{1}{2} \\ \frac{1}{2}, & \frac{1}{2} \le s < 1. \end{cases}$$

Solution.
(a) X can only take the values 0, 1, ..., 10, so X is a discrete random variable.
(b) Y can take any positive real value, so Y is a continuous random variable.
(c) Z is a mixed random variable since Z is continuous in the interval $(0, \frac{1}{2})$ and discrete on the interval $[\frac{1}{2}, 1)$ ■

Example 13.2
The sample space of the experiment of tossing a coin 3 times is given by

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

Let X = # of Heads in 3 tosses. Find the range of X.

Solution.
We have

$$\begin{array}{lllllllllll} X(HHH) & = & 3 & X(HHT) & = & 2 & X(HTH) & = & 2 & X(HTT) & = & 1 \\ X(THH) & = & 2 & X(THT) & = & 1 & X(TTH) & = & 1 & X(TTT) & = & 0 \end{array}$$

Thus, the range of X consists of $\{0, 1, 2, 3\}$ so that X is a discrete random variable ■

We use upper-case letters $X, Y, Z,$ etc. to represent random variables. We use small letters $x, y, z,$ etc to represent possible values that the corresponding random variables $X, Y, Z,$ etc. can take. The statement $X = x$ defines an event consisting of all outcomes with X-measurement equal to x which is the set $\{s \in S : X(s) = x\}$. For instance, considering the random variable of the previous example, the statement "X = 2" is the event $\{HHT, HTH, THH\}$. Because the value of a random variable is determined by the outcomes of the experiment, we may assign probabilities to the possible values of the random variable. For example, $\Pr(X = 2) = \frac{3}{8}$.

Example 13.3
Consider the experiment consisting of 2 rolls of a fair 4-sided die. Let X be a random variable, equal to the maximum of the 2 rolls. Complete the following table

x	1	2	3	4
Pr(X=x)				

Solution.
The sample space of this experiment is

$$S = \{(1,1),(1,2),(1,3),(1,4),(2,1),(2,2),(2,3),(2,4),(3,1),(3,2),(3,3),\\ (3,4),(4,1),(4,2),(4,3),(4,4)\}.$$

Thus,

x	1	2	3	4
Pr(X=x)	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{7}{16}$

Example 13.4
A class consisting of five male students and five female students has taken the GRE examination. All ten students got different scores on the test. The students are ranked according to their scores on the test. Assume that all possible rankings are equally likely. Let X denote the highest ranking achieved by a male student. Find $\Pr(X = i), i = 1, 2, \cdots, 10$.

Solution.
Since 6 is the lowest possible rank attainable by the highest-scoring male, we must have $\Pr(X = 7) = \Pr(X = 8) = \Pr(X = 9) = \Pr(X = 10) = 0$.
For $X = 1$ (male is highest-ranking scorer), we have 5 possible choices out of 10 for the top spot that satisfy this requirement; hence

$$\Pr(X = 1) = \frac{5 \cdot 9!}{10!} = \frac{1}{2}.$$

For $X = 2$ (male is 2nd-highest scorer), we have 5 possible choices for the top female, then 5 possible choices for the male who ranked 2nd overall, and then any arrangement of the remaining 8 individuals is acceptable (out of 10! possible arrangements of 10 individuals); hence,

$$\Pr(X = 2) = \frac{5 \cdot 5 \cdot 8!}{10!} = \frac{5}{18}.$$

For $X = 3$ (male is 3rd-highest scorer), acceptable configurations yield (5)(4)=20 possible choices for the top 2 females, 5 possible choices for the male who ranked 3rd overall, and 7! different arrangement of the remaining

7 individuals (out of a total of 10! possible arrangements of 10 individuals); hence,

$$\Pr(X = 3) = \frac{5 \cdot 4 \cdot 5 \cdot 7!}{10!} = \frac{5}{36}.$$

Similarly, we have

$$\begin{aligned}
\Pr(X = 4) =& \frac{5 \cdot 4 \cdot 3 \cdot 5 \cdot 6!}{10!} = \frac{5}{84} \\
\Pr(X = 5) =& \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 5 \cdot 5!}{10!} = \frac{5}{252} \\
\Pr(X = 6) =& \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 5 \cdot 4!}{10!} = \frac{1}{252} \blacksquare
\end{aligned}$$

Practice Problems

Problem 13.1
Determine whether the random variable is discrete, continuous or mixed.
(a) X is a randomly selected number in the interval $(0, 1)$. (b) Y is the number of heart beats per minute.
(c) Z is the number of calls at a switchboard in a day.
(d) $U : (0, 1) \to \mathbb{R}$ defined by $U(s) = 2s - 1$.
(e) $V : (0, 1) \to \mathbb{R}$ defined by $V(s) = 2s - 1$ for $0 < s < \frac{1}{2}$ and $V(s) = 1$ for $\frac{1}{2} \leq s < 1$.

Problem 13.2
Two apples are selected at random and removed in succession and without replacement from a bag containing five golden apples and three red apples. List the elements of the sample space, the corresponding probabilities, and the corresponding values of the random variable X, where X is the number of golden apples selected.

Problem 13.3
Suppose that two fair dice are rolled so that the sample space is $S = \{(i, j) : 1 \leq i, j \leq 6\}$. Let X be the random variable $X(i, j) = i + j$. Find $\Pr(X = 6)$.

Problem 13.4
Let X be a random variable with probability distribution table given below

x	0	10	20	50	100
Pr(X=x)	0.4	0.3	0.15	0.1	0.05

Find $\Pr(X < 50)$.

Problem 13.5
You toss a coin repeatedly until you get heads. Let X be the random variable representing the number of times the coin flips until the first head appears. Find $\Pr(X = n)$ where n is a positive integer.

Problem 13.6
A couple is expecting the arrival of a new boy. They are deciding on a name from the list $S = \{$ Steve, Stanley, Joseph, Elija $\}$. Let $X(\omega) =$ first letter in name. Find $\Pr(X = S)$.

Problem 13.7 ‡

The number of injury claims per month is modeled by a random variable N with

$$\Pr(N = n) = \frac{1}{(n+1)(n+2)}, \quad n \geq 0.$$

Determine the probability of at least one claim during a particular month, given that there have been at most four claims during that month.

Problem 13.8

Let X be a discrete random variable with the following probability table

x	1	5	10	50	100
Pr(X=x)	0.02	0.41	0.21	0.08	0.28

Compute $\Pr(X > 4 | X \leq 50)$.

Problem 13.9

Shooting is one of the sports listed in the Olympic games. A contestant shoots three times, independently. The probability of hiting the target in the first try is 0.7, in the second try 0.5, and in the third try 0.4. Let X be the discrete random variable representing the number of successful shots among these three.

(a) Find a formula for the piecewise defined function $X : \Omega \to \mathbb{R}$.

(b) Find the event corresponding to $X = 0$. What is the probability that he misses all three shots; i.e., $\Pr(X = 0)$?

(c) What is the probability that he succeeds exactly once among these three shots; i.e $\Pr(X = 1)$?

(d) What is the probability that he succeeds exactly twice among these three shots; i.e $\Pr(X = 2)$?

(e) What is the probability that he makes all three shots; i.e $\Pr(X = 3)$?

Problem 13.10

Let X be a discrete random variable with range $\{0, 1, 2, 3, \cdots\}$. Suppose that

$$\Pr(X = 0) = \Pr(X = 1), \quad \Pr(X = k + 1) = \frac{1}{k}\Pr(X = k), \quad k = 1, 2, 3, \cdots$$

Find $\Pr(0)$.

Problem 13.11 ‡
Under an insurance policy, a maximum of five claims may be filed per year by a policyholder. Let p_n be the probability that a policyholder files n claims during a given year, where $n = 0, 1, 2, 3, 4, 5$. An actuary makes the following observations:
(i) $p_n \geq p_{n+1}$ for $0 \leq n \leq 4$
(ii) The difference between p_n and p_{n+1} is the same for $0 \leq n \leq 4$
(iii) Exactly 40% of policyholders file fewer than two claims during a given year.
Calculate the probability that a random policyholder will file more than three claims during a given year.

14 Probability Mass Function and Cumulative Distribution Function

For a discrete random variable X, we define the **probability distribution** or the **probability mass function**(abbreviated pmf) by the equation

$$p(x) = \Pr(X = x).$$

That is, a probability mass function gives the probability that a discrete random variable is exactly equal to some value.
The pmf can be an equation, a table, or a graph that shows how probability is assigned to possible values of the random variable.

Example 14.1
Suppose a variable X can take the values 1, 2, 3, or 4. The probabilities associated with each outcome are described by the following table:

x	1	2	3	4
$p(x)$	0.1	0.3	0.4	0.2

Draw the probability histogram.

Solution.
The probability histogram is shown in Figure 14.1 ■

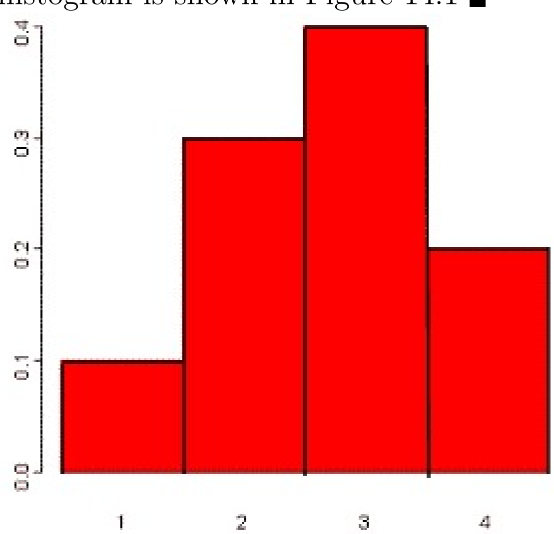


Figure 14.1

Example 14.2
A committee of 4 is to be selected from a group consisting of 5 men and 5 women. Let X be the random variable that represents the number of women in the committee. Create the probability mass distribution.

Solution.
For $x = 0, 1, 2, 3, 4$ we have

$$p(x) = \frac{\left(\begin{array}{c} 5 \\ x \end{array}\right)\left(\begin{array}{c} 5 \\ 4-x \end{array}\right)}{\left(\begin{array}{c} 10 \\ 4 \end{array}\right)}.$$

The probability mass function can be described by the table

x	0	1	2	3	4
$p(x)$	$\frac{5}{210}$	$\frac{50}{210}$	$\frac{100}{210}$	$\frac{50}{210}$	$\frac{5}{210}$

Example 14.3
Consider the experiment of rolling a fair die twice. Let $X(i, j) = \max\{i, j\}$. Find the equation of $\Pr(x)$.

Solution.
The pmf of X is

$$\begin{aligned} p(x) =& \left\{ \begin{array}{cc} \frac{2x-1}{36} & \text{if } x = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise} \end{array} \right. \\ =& \frac{2x-1}{36} I_{\{1,2,3,4,5,6\}}(x) \end{aligned}$$

where

$$I_{\{1,2,3,4,5,6\}}(x) = \left\{ \begin{array}{ll} 1 & \text{if } x \in \{1, 2, 3, 4, 5, 6\} \\ 0 & \text{otherwise} \ \blacksquare \end{array} \right.$$

In general, we define the **indicator function** of a set A to be the function

$$I_A(x) = \left\{ \begin{array}{ll} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{array} \right.$$

Note that if the range of a random variable is $\Omega = \{x_1, x_2, \cdots\}$ then

$$p(x) \geq 0, \ x \in \Omega$$

and
$$\sum_{x\in\Omega} p(x) = 1.$$
All random variables (discrete, continuous or mixed) have a **distribution function** or a **cumulative distribution function**, abbreviated cdf. It is a function giving the probability that the random variable X is less than or equal to x, for every value x. For a discrete random variable, the cumulative distribution function is found by summing up the probabilities. That is,
$$F(a) = \Pr(X \le a) = \sum_{x\le a} p(x).$$

Example 14.4
Given the following pmf
$$p(x) = \begin{cases} 1, & \text{if } x = a \\ 0, & \text{otherwise} \end{cases}$$
Find a formula for $F(x)$ and sketch its graph.

Solution.
A formula for $F(x)$ is given by
$$F(x) = \begin{cases} 0, & \text{if } x < a \\ 1, & \text{otherwise} \end{cases}$$
Its graph is given in Figure 14.2 ■

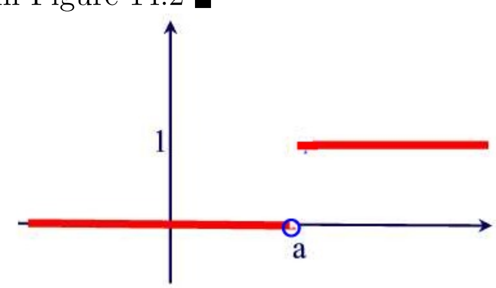


Figure 14.2

For discrete random variables the cumulative distribution function will always be a step function with jumps at each value of x that has probability greater than 0. Note the value of $F(x)$ is assigned to the top of the jump.

Example 14.5
Consider the following probability mass distribution

x	1	2	3	4
$p(x)$	0.25	0.5	0.125	0.125

Find a formula for $F(x)$ and sketch its graph.

Solution.
The cdf is given by

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.25 & 1 \leq x < 2 \\ 0.75 & 2 \leq x < 3 \\ 0.875 & 3 \leq x < 4 \\ 1 & 4 \leq x. \end{cases}$$

Its graph is given in Figure 14.3 ■

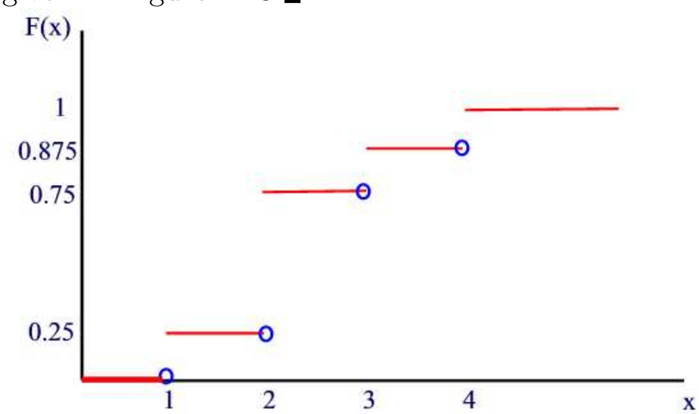


Figure 14.3

Note that the size of the step at any of the values 1,2,3,4 is equal to the probability that X assumes that particular value. That is, we have

Theorem 14.1
If the range of a discrete random variable X consists of the values $x_1 < x_2 < \cdots < x_n$ then $p(x_1) = F(x_1)$ and

$$p(x_i) = F(x_i) - F(x_{i-1}), \quad i = 2, 3, \cdots, n$$

Proof.
Because $F(x) = 0$ for $x < x_1$ then $F(x_1) = \Pr(X \leq x_1) = \Pr(X < x_1) + \Pr(X = x_1) = \Pr(X = x_1) = p(x_1)$. Now, for $i = 2, 3, \cdots, n$, let $A = \{s \in S : X(s) > x_{i-1}\}$ and $B = \{s \in S : X(s) \leq x_i\}$. Thus, $A \cup B = S$. We have

$$\begin{aligned}
\Pr(x_{i-1} < X \leq x_i) =& \Pr(A \cap B)\\
=& \Pr(A) + \Pr(B) - \Pr(A \cup B)\\
=& 1 - F(x_{i-1}) + F(x_i) - 1\\
=& F(x_i) - F(x_{i-1}) \blacksquare
\end{aligned}$$

Example 14.6
If the cumulative distribution function of X is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{16} & 0 \leq x < 1 \\ \frac{5}{16} & 1 \leq x < 2 \\ \frac{11}{16} & 2 \leq x < 3 \\ \frac{15}{16} & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

find the pmf of X.

Solution.
Making use of the previous theorem, we get $p(0) = \frac{1}{16}, p(1) = \frac{1}{4}, p(2) = \frac{3}{8}, p(3) = \frac{1}{4}$, and $p(4) = \frac{1}{16}$ and 0 otherwise $\blacksquare$

Practice Problems

Problem 14.1
Consider the experiment of tossing a fair coin three times. Let X denote the random variable representing the total number of heads.
(a) Describe the probability mass function by a table.
(b) Describe the probability mass function by a histogram.

Problem 14.2
In the previous problem, describe the cumulative distribution function by a formula and by a graph.

Problem 14.3
Toss a pair of fair dice. Let X denote the sum of the dots on the two faces. Find the probability mass function.

Problem 14.4
A box of six apples has one roten apple. Randomly draw one apple from the box, without replacement, until the roten apple is found. Let X denote the number of apples drawn until the roten apple is found. Find the probability mass function of X and draw its histogram.

Problem 14.5
In the experiment of rolling two dice, let X be the random variable representing the number of even numbers that appear. Find the probability mass function of X.

Problem 14.6
Let X be a random variable with pmf

$$p(n) = \frac{1}{3}\left(\frac{2}{3}\right)^n, \quad n = 0, 1, 2, \cdots .$$

Find a formula for $F(n)$.

Problem 14.7
A box contains 100 computer mice of which 95 are defective.
(a) One mouse is taken from the box at a time (without replacement) until a nondefective mouse is found. Let X be the number of mouses you have

to take out in order to find one that is not defective. Find the probability distribution of X.
(b) Exactly 10 mouses were taken from the box and then each of the 10 mouses is tested. Let Y denote the number of nondefective mouses among the 10 that were taken out. Find the probability distribution of Y.

Problem 14.8
Let X be a discrete random variable with cdf given by

$$F(x) = \begin{cases} 0 & x < -4 \\ \frac{3}{10} & -4 \leq x < 1 \\ \frac{7}{10} & 1 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

Find a formula of $p(x)$.

Problem 14.9
A game consists of randomly selecting two balls without replacement from an urn containing 3 red balls and 4 blue balls. If the two selected balls are of the same color then you win \$2. If they are of different colors then you lose \$1. Let X denote your gain/lost. Find the probability mass function of X.

Problem 14.10
An unfair coin is tossed three times. The probability of tails on any particular toss is known to be $\frac{2}{3}$. Let X denote the number of heads.
(a) Find the probability mass function.
(b) Graph the cumulative distribution function for X.

Problem 14.11
A lottery game consists of matching three numbers drawn (without replacement) from a set of 15 numbers. Let X denote the random variable representing the numbers on your tickets that match the winning numbers. Find the cumulative distribution of X.

Problem 14.12
If the cumulative distribution function of X is given by

$$F(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{36} & 2 \le x < 3 \\ \frac{3}{36} & 3 \le x < 4 \\ \frac{6}{36} & 4 \le x < 5 \\ \frac{10}{36} & 5 \le x < 6 \\ \frac{15}{36} & 6 \le x < 7 \\ \frac{21}{36} & 7 \le x < 8 \\ \frac{26}{36} & 8 \le x < 9 \\ \frac{30}{36} & 9 \le x < 10 \\ \frac{33}{36} & 10 \le x < 11 \\ \frac{35}{36} & 11 \le x < 12 \\ 1 & x \ge 12 \end{cases}$$

find the probability distribution of X.

15 Expected Value of a Discrete Random Variable

A cube has three red faces, two green faces, and one blue face. A game consists of rolling the cube twice. You pay \$2 to play. If both faces are the same color, you are paid \$5(that is you win \$3). If not, you lose the \$2 it costs to play. Will you win money in the long run? Let W denote the event that you win. Then $W = \{RR, GG, BB\}$ and

$$\Pr(W) = \Pr(RR) + \Pr(GG) + \Pr(BB) = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{3}\cdot\frac{1}{3} + \frac{1}{6}\cdot\frac{1}{6} = \frac{7}{18} \approx 39\%.$$

Thus, $\Pr(L) = \frac{11}{18} = 61\%$. Hence, if you play the game 18 times you expect to win 7 times and lose 11 times on average. So your winnings in dollars will be $3 \times 7 - 2 \times 11 = -1$. That is, you can expect to lose \$1 if you play the game 18 times. On the average, you will lose \$ $\frac{1}{18}$ per game (about 6 cents). This can be found also using the equation

$$3 \times \frac{7}{18} - 2 \times \frac{11}{18} = -\frac{1}{18}$$

If we let X denote the winnings of this game then the range of X consists of the two numbers 3 and -2 which occur with respective probability 0.39 and 0.61. Thus, we can write

$$E(X) = 3 \times \frac{7}{18} - 2 \times \frac{11}{18} = -\frac{1}{18}.$$

We call this number the expected value of X. More formally, let the range of a discrete random variable X be a sequence of numbers $x_1, x_2, \cdots, x_k$, and let $\Pr(x)$ be the corresponding probability mass function. Then the **expected value** of X is

$$E(X) = x_1 p(x_1) + x_2 p(x_2) + \cdots + x_k p(x_k).$$

The following is a justification of the above formula. Suppose that X has k possible values $x_1, x_2, \cdots, x_k$ and that

$$p_i = \Pr(X = x_i) = p(x_i), i = 1, 2, \cdots, k.$$

Suppose that in n repetitions of the experiment, the number of times that X takes the value x_i is n_i. Then the sum of the values of X over the n repetitions is

$$n_1 x_1 + n_2 x_2 + \cdots + n_k x_k$$

and the average value of X is

$$\frac{n_1x_1 + n_2x_2 + \cdots + n_kx_k}{n} = \frac{n_1}{n}x_1 + \frac{n_2}{n}x_2 + \cdots + \frac{n_k}{n}x_k.$$

But $\Pr(X = x_i) = \lim_{n\to\infty} \frac{n_i}{n}$. Thus, the average value of X approaches

$$E(X) = x_1p(x_1) + x_2p(x_2) + \cdots + x_kp(x_k).$$

The expected value of X is also known as the **mean** value.

Example 15.1 ‡
Suppose that an insurance company has broken down yearly automobile claims for drivers from age 16 through 21 as shown in the following table.

Amount of claim	Probability
\$ 0	0.80
\$ 2000	0.10
\$ 4000	0.05
\$ 6000	0.03
\$ 8000	0.01
\$ 10000	0.01

How much should the company charge as its average premium in order to break even on costs for claims?

Solution.
Let X be the random variable of the amount of claim. Finding the expected value of X we have

$$E(X) = 0(.80)+2000(.10)+4000(.05)+6000(.03)+8000(.01)+10000(.01) = 760$$

Since the average claim value is \$760, the average automobile insurance premium should be set at \$760 per year for the insurance company to break even ■

Example 15.2
Let A be a nonempty set. Consider the random variable I with range 0 and 1 and with pmf the indicator function I_A where

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases}$$

Find $E(I)$.

Solution.
Since $\Pr(1) = \Pr(A)$ and $\Pr(0) = \Pr(A^c)$, we have

$$E(I) = 1 \cdot \Pr(A) + 0 \cdot \Pr(A^c) = \Pr(A)$$

That is, the expected value of I is just the probability of A ■

Example 15.3
An insurance policy provides the policyholder with a payment of \$1,000 if a death occurs within 5 years. Let X be the random variable of the amount paid by an insurance company to the policyholder. Suppose that the probability of death of the policyholder within 5 years is estimated to be 0.15.
(a) Find the probability distribution of X.
(b) What is the most the policyholder should be willing to pay for this policy?

Solution.
(a) $\Pr(X = 1,000) = 0.15$ and $\Pr(X = 0) = 0.85$.
(b) $E(X) = 1000 \times 0.15 + 0 \times 0.85 = 150$. Thus, the policyholder expected payout is \$150, so he/she should not be willing to pay more than \$150 for the policy ■

Example 15.4
You have a fancy car video system in your car and you feel you want to insure it against theft. An insurance company offers you a \$2000 1-year coverage for a premium of \$225. The probability that the theft will occur is 0.1. What is your expected return from this policy?

Solution.
Let X be the random variable of the profit/loss from this policy to policyholder. Then either $X = 1,775$ with probability 0.1 or $X = -225$ with porbability 0.9. Thus, the expected return of this policy is

$$E(X) = 1,775(0.1) + (-225)(0.9) = -\$25.$$

That is, by insuring the car video system for many years, and under the same circumstances, you will expect a net loss of \$25 per year to the insurance company ■

Remark 15.1
The expected value (or mean) is related to the physical property of center of mass. If we have a weightless rod in which weights of mass $\Pr(x)$ located at a distance x from the left endpoint of the rod then the point at which the rod is balanced is called the **center of mass.** If α is the center of mass then we must have $\sum_x (x - \alpha)p(x) = 0$. This equation implies that $\alpha = \sum_x xp(x) = E(X)$. Thus, the expected value tells us something about the center of the probability mass function.

Practice Problems

Problem 15.1
Consider the experiment of rolling two dice. Let X be the random variable representing the sum of the two faces. Find $E(X)$.

Problem 15.2
A game consists of rolling two dice. The sum of the two faces is a positive integer between 2 and 12. For each such a value, you win an amount of money as shown in the table below.

Score	2	3	4	5	6	7	8	9	10	11	12
$ won	4	6	8	10	20	40	20	10	8	6	4

Compute the expected value of this game.

Problem 15.3
A game consists of rolling two dice. The game costs $2 to play. If a sum of 7 appears you win $10 otherwise you lose your $2. Would you be making money, losing money, or coming out about even if you keep playing this game? Explain.

Problem 15.4
A game consists of rolling two dice. The game costs $8 to play. You get paid the sum of the numbers in dollars that appear on the dice. What is the expected value of this game (long-run average gain or loss per game)?

Problem 15.5
A storage company provides insurance coverage for items stored on its premises. For items valued at $800, the probability that $400 worth of items of being stolen is 0.01 while the probability the whole items being stolen is 0.0025. Assume that these are the only possible kinds of expected loss. How much should the storage company charge for people with this coverage in order to cover the money they pay out and to make an additional $20 profit per person on the average?

Problem 15.6
A game consists of spinning a spinner with payoff as shown in Figure 15.1. The cost of playing is $2 per spin. What is the expected return to the owner

of this game?

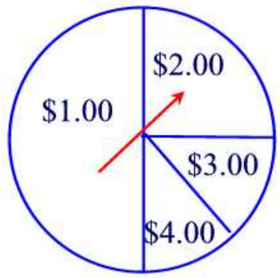


Figure 15.1

Problem 15.7
Consider a game that costs \$1 to play. The probability of losing is 0.7. The probability of winning \$50 is 0.1, and the probability of winning \$35 is 0.2. Would you expect to win or lose if you play this game 10 times?

Problem 15.8
A lottery type game consists of matching the correct three numbers that are selected from the numbers 1 through 12. The cost of one ticket is \$1. If your ticket matched the three selected numbers, you win \$100. What are your expected earnings?

Problem 15.9 ‡
Two life insurance policies, each with a death benefit of 10,000 and a one-time premium of 500, are sold to a couple, one for each person. The policies will expire at the end of the tenth year. The probability that only the wife will survive at least ten years is 0.025, the probability that only the husband will survive at least ten years is 0.01, and the probability that both of them will survive at least ten years is 0.96 .
What is the expected excess of premiums over claims, given that the husband survives at least ten years?

Problem 15.10
An urn contains 30 marbles of which 8 are black, 12 are red, and 10 are blue. Randomly, select four marbles without replacement. Let X be the number black marbles in the sample of four.
(a) What is the probability that no black marble was selected?
(b) What is the probability that exactly one black marble was selected?
(c) Compute $E(X)$.

Problem 15.11
The distribution function of a discrete random variable X is given by

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.2 & -2 \leq x < 0 \\ 0.5 & 0 \leq x < 2.2 \\ 0.6 & 2.2 \leq x < 3 \\ 0.6 + q & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

Suppose that $\Pr(X > 3) = 0.1$.
(a) Determine the value of q?
(b) Compute $\Pr(X^2 > 2)$.
(c) Find $p(0),\ p(1)$? and $p(\Pr(X \leq 0))$.
(d) Find the formula of the probability mass function $p(x)$.
(e) Compute $E(X)$.

Problem 15.12
A computer store specializes in selling used laptops. The laptops can be classified as either in good condition or in fair condition. Assume that the store salesperson is able to tell whether a laptop is in good or fair condition. However, a buyer in the store can not tell the difference. Suppose that buyers are aware that the probability of a laptop of being in good condition is 0.4. A laptop in good condition costs the store \$400 and a buyer is willing to pay \$525 for it whereas a laptop in fair conidition costs the store \$200 and a buyer is willing to pay for \$300 for it.
(a) Find the expected value of a used laptop to a buyer who has no extra information.
(b) Assuming that buyers will not pay more than their expected value for a used laptop, will sellers ever sell laptops in good condition?

Problem 15.13
An urn contains 10 marbles in which 3 are black. Four of the marbles are selected at random (without replacement) and are tested for the black color. Define the random variable X to be the number of the selected marbles that are not black.
(a) Find the probability mass function of X.
(b) What is the cumulative distribution function of X?
(c) Find the expected value of X.

Problem 15.14 ‡
An auto insurance company is implementing a new bonus system. In each month, if a policyholder does not have an accident, he or she will receive a 5.00 cash-back bonus from the insurer.
Among the 1,000 policyholders of the auto insurance company, 400 are classified as low-risk drivers and 600 are classified as high-risk drivers.
In each month, the probability of zero accidents for high-risk drivers is 0.80 and the probability of zero accidents for low-risk drivers is 0.90.
Calculate the expected bonus payment from the insurer to the 1000 policyholders in one year.

16 Expected Value of a Function of a Discrete Random Variable

If we apply a function $g(\cdot)$ to a random variable X, the result is another random variable $Y = g(X)$. For example, $X^2, \log X, \frac{1}{X}$ are all random variables derived from the original random variable X.
In this section we are interested in finding the expected value of this new random variable. But first we look at an example.

Example 16.1
Let X be a discrete random variable with range $\{-1, 0, 1\}$ and probabilities $\Pr(X = -1) = 0.2, \Pr(X = 0) = 0.5$, and $\Pr(X = 1) = 0.3$. Compute $E(X^2)$.

Solution.
Let $Y = X^2$. Then the range of Y is $\{0, 1\}$. Also, $\Pr(Y = 0) = \Pr(X = 0) = 0.5$ and $\Pr(Y = 1) = \Pr(X = -1) + \Pr(X = 1) = 0.2 + 0.3 = 0.5$ Thus, $E(X^2) = 0(0.5) + 1(0.5) = 0.5$. Note that $E(X) = -1(0.2) + 0(0.5) + 1(0.3) = 0.1$ so that $E(X^2) \neq (E(X))^2$ ■

Now, if X is a discrete random variable and $g(x) = x$ then we know that

$$E(g(X)) = E(X) = \sum_{x \in D} xp(x)$$

where D is the range of X and $p(x)$ is its probability mass function. This suggests the following result for finding $E(g(X))$.

Theorem 16.1
If X is a discrete random variable with range D and pmf $\Pr(x)$, then the expected value of any function $g(X)$ is computed by

$$E(g(X)) = \sum_{x \in D} g(x)p(x).$$

Proof.
Let D be the range of X and D' be the range of $g(X)$. Thus,

$$D' = \{g(x) : x \in D\}.$$

For each $y \in D'$ we let $A_y = \{x \in D : g(x) = y\}$. We will show that $\{s \in S : g(X)(s) = y\} = \bigcup_{x \in A_y} \{s \in S : X(s) = x\}$. The prove is by double inclusions. Let $s \in S$ be such that $g(X)(s) = y$. Then $g(X(s)) = y$. Since $X(s) \in D$, there is an $x \in D$ such that $x = X(s)$ and $g(x) = y$. This shows that $s \in \cup_{x \in A_y} \{s \in S : X(s) = x\}$. For the converse, let $s \in \cup_{x \in A_y} \{s \in S : X(s) = x\}$. Then there exists $x \in D$ such that $g(x) = y$ and $X(s) = x$. Hence, $g(X)(s) = g(x) = y$ and this implies that $s \in \{s \in S : g(X)(s) = y\}$. Next, if x_1 and x_2 are two distinct elements of A_y and $w \in \{s \in S : X(s) = x_1\} \cap \{t \in S : X(t) = x_2\}$ then this leads to $x_1 = x_2$, a contradiction. Hence, $\{s \in S : X(s) = x_1\} \cap \{t \in S : X(t) = x_2\} = \emptyset$.
From the above discussion we are in a position to find $p_Y(y)$, the pmf of $Y = g(X)$, in terms of the pmf of X. Indeed,

$$\begin{aligned}
p_Y(y) =& \Pr(Y = y) \\
=& \Pr(g(X) = y) \\
=& \sum_{x \in A_y} \Pr(X = x) \\
=& \sum_{x \in A_y} p(x)
\end{aligned}$$

Now, from the definition of the expected value we have

$$\begin{aligned}
E(g(X)) =& E(Y) = \sum_{y \in D'} y p_Y(y) \\
=& \sum_{y \in D'} y \sum_{x \in A_y} p(x) \\
=& \sum_{y \in D'} \sum_{x \in A_y} g(x) p(x) \\
=& \sum_{x \in D} g(x) p(x)
\end{aligned}$$

Note that the last equality follows from the fact that D is the disjoint unions of the A_y ■

Example 16.2
Let X be the number of points on the side that comes up when rolling a fair die. Find the expected value of $g(X) = 2X^2 + 1$.

Solution.
Since each possible outcome has the probability $\frac{1}{6}$, we get

$$\begin{aligned}E[g(X)] =&\sum_{i=1}^{6}(2i^2+1)\cdot\frac{1}{6}\\ =&\frac{1}{6}\left(6+2\sum_{i=1}^{6}i^2\right)\\ =&\frac{1}{6}\left(6+2\frac{6(6+1)(2\cdot 6+1)}{6}\right)=\frac{94}{3}\ \blacksquare\end{aligned}$$

As a consequence of the above theorem we have the following result.

Corollary 16.1
If X is a discrete random variable, then for any constants a and b we have

$$E(aX+b)=aE(X)+b.$$

Proof.
Let D denote the range of X. Then

$$\begin{aligned}E(aX+b) =&\sum_{x\in D}(ax+b)p(x)\\ =&a\sum_{x\in D}xp(x)+b\sum_{x\in D}p(x)\\ =&aE(X)+b\ \blacksquare\end{aligned}$$

A similar argument establishes

$$E(aX^2+bX+c)=aE(X^2)+bE(X)+c.$$

Example 16.3
Let X be a random variable with $E(X)=6$ and $E(X^2)=45$, and let $Y=20-2X$. Find $E(Y)$ and $E(Y^2)-[E(Y)]^2$.

Solution.
By the properties of expectation,

$$\begin{aligned}E(Y) =&E(20-2X) = 20-2E(X) = 20-12 = 8\\E(Y^2) =&E(400-80X+4X^2) = 400-80E(X)+4E(X^2) = 100\end{aligned}$$

$$E(Y^2)-(E(Y))^2 = 100-64 = 36 \blacksquare$$

We conclude this section with the following definition. If $g(x) = x^n$ then we call $E(X^n) = \sum_x x^n p(x)$ the n^{th} **moment about the origin** of X or the n^{th} **raw moment**. Thus, $E(X)$ is the first moment of X.

Example 16.4
Show that $E(X^2) = E(X(X-1)) + E(X)$.

Solution.
Let D be the range of X. We have

$$\begin{aligned}E(X^2) =&\sum_{x\in D} x^2 p(x)\\=&\sum_{x\in D}(x(x-1)+x)p(x)\\=&\sum_{x\in D} x(x-1)p(x) + \sum_{x\in D} xp(x) = E(X(X-1)) + E(X) \blacksquare\end{aligned}$$

Remark 16.1
In our definition of expectation the set D can be countably infinite. It is possible to have a random variable with undefined expectation as seen in the next example.

Example 16.5
If the probability distribution of X is given by

$$p(x) = \left(\frac{1}{2}\right)^x, \quad x = 1, 2, 3, \cdots$$

show that $E(2^X)$ does not exist.

Solution.
We have

$$E(2^X) = (2^1)\frac{1}{2^1} + (2^2)\frac{1}{2^2} + \cdots = \sum_{n=1}^{\infty} 1$$

The series on the right is divergent so that $E(2^X)$ does not exist $\blacksquare$

Practice Problems

Problem 16.1
Suppose that X is a discrete random variable with probability mass function

$$p(x) = cx^2, \quad x = 1, 2, 3, 4.$$

(a) Find the value of c.
(b) Find $E(X)$.
(c) Find $E(X(X-1))$.

Problem 16.2
A random variable X has the following probability mass function defined in tabular form

x	-1	1	2
$p(x)$	2c	3c	4c

(a) Find the value of c.
(b) Compute $p(-1), p(1)$, and $p(2)$.
(c) Find $E(X)$ and $E(X^2)$.

Problem 16.3
Let X be a random variable with range $\{1, 2, 3, 4, 5, 6\}$. Suppose that $p(x) = kx$ for some positive constant k.
(a) Determine the value of k.
(b) Find $\Pr(X = x)$ for x even.
(c) Find the expected value of X.

Problem 16.4
Let X be a discrete random variable. Show that $E(aX^2+bX+c) = aE(X^2)+bE(X)+c$.

Problem 16.5
Consider a random variable X whose probability mass function is given by

$$p(x) = \begin{cases} 0.1 & x = -3 \\ 0.2 & x = 0 \\ 0.3 & x = 2.2 \\ 0.1 & x = 3 \\ 0.3 & x = 4 \\ 0 & \text{otherwise} \end{cases}$$

Let $F(x)$ be the corresponding cdf. Find $E(F(X))$.

Problem 16.6 ‡
An insurance policy pays 100 per day for up to 3 days of hospitalization and 50 per day for each day of hospitalization thereafter.
The number of days of hospitalization, X, is a discrete random variable with probability function

$$p(k) = \begin{cases} \frac{6-k}{15} & k = 1, 2, 3, 4, 5 \\ 0 & \text{otherwise} \end{cases}$$

Determine the expected payment for hospitalization under this policy.

Problem 16.7 ‡
An insurance company sells a one-year automobile policy with a deductible of 2 . The probability that the insured will incur a loss is 0.05 . If there is a loss, the probability of a loss of amount N is $\frac{K}{N}$, for $N = 1, \cdots, 5$ and K a constant. These are the only possible loss amounts and no more than one loss can occur.
Determine the net premium for this policy.

Problem 16.8
Consider a random variable X whose probability mass function is given by

$$p(x) = \begin{cases} 0.2 & x = -1 \\ 0.3 & x = 0 \\ 0.1 & x = 0.2 \\ 0.1 & x = 0.5 \\ 0.3 & x = 4 \\ 0 & \text{otherwise} \end{cases}$$

Find $E(p(x))$.

Problem 16.9
A box contains 7 marbles of which 3 are red and 4 are blue. Randomly select two marbles without replacement. If the marbles are of the same color then you win \$2, otherwise you lose \$1. Let X be the random variable representing your net winnings.
(a) Find the probability mass function of X.
(b) Compute $E(2^X)$.

Problem 16.10

Three kinds of tickets are sold at a movie theater: children (for \$3), adult (for \$8), and seniors (for \$5). Let C denote the number of children tickets sold, A number of adult tickets, and S number of senior tickets. You are given: $E[C] = 45,\ E[A] = 137,\ E[S] = 34$. Assume the number of tickets sold is indepndent.

Any particular movie costs \$300 to show, regardless of the audience size.

(a) Write a formula relating $C, A,$ and S to the theater's profit P for a particular movie.

(b) Find $E(P)$.

17 Variance and Standard Deviation

In the previous section we learned how to find the expected values of various functions of random variables. The most important of these are the variance and the standard deviation which give an idea about how spread out the probability mass function is about its expected value.
The expected squared distance between the random variable and its mean is called the **variance** of the random variable. The positive square root of the variance is called the **standard deviation** of the random variable. If σ_X denotes the standard deviation then the variance is given by the formula

$$\mathrm{Var(X)} = \sigma_X^2 = E\left[(X - E(X))^2\right]$$

The variance of a random variable is typically calculated using the following formula

$$\begin{aligned}\mathrm{Var}(X) =&E[(X - E(X))^2]\\=&E[X^2 - 2XE(X) + (E(X))^2]\\=&E(X^2) - 2E(X)E(X) + (E(X))^2\\=&E(X^2) - (E(X))^2\end{aligned}$$

where we have used the result of Problem 16.4.

Example 17.1
Find the variance of the random variable X with probability distribution $\Pr(X = 1) = \Pr(X = -1) = \frac{1}{2}$.

Solution.
Since $E(X) = 1 \times \frac{1}{2} - 1 \times \frac{1}{2} = 0$ and $E(X^2) = 1^2\frac{1}{2} + (-1)^2 \times \frac{1}{2} = 1$ we find $\mathrm{Var}(X) = 1 - 0 = 1$ ■

A useful identity is given in the following result

Theorem 17.1
If X is a discrete random variable then for any constants a and b we have

$$\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$$

Proof.
Since $E(aX+b)=aE(X)+b$, we have

$$\begin{aligned}\text{Var}(aX+b)=&E\left[(aX+b-E(aX+b))^2\right]\\=&E[a^2(X-E(X))^2]\\=&a^2E((X-E(X))^2)\\=&a^2\text{Var}(X)\ \blacksquare\end{aligned}$$

Remark 17.1
Note that the units of Var(X) is the square of the units of X. This motivates the definition of the standard deviation $\sigma_X=\sqrt{\text{Var}(X)}$ which is measured in the same units as X.

Example 17.2
In a recent study, it was found that tickets cost to the Dallas Cowboys football games averages \$80 with a variance of 105 square dollar. What will be the variance of the cost of tickets if 3% tax is charged on all tickets?

Solution.
Let X be the current ticket price and Y be the new ticket price. Then $Y=1.03X$. Hence,

$$\text{Var}(Y)=\text{Var}(1.03X)=1.03^2\text{Var}(X)=(1.03)^2(105)=111.3945\ \blacksquare$$

Example 17.3
In the experiment of rolling one die, let X be the number on the face that comes up. Find the variance and standard deviation of X.

Solution.
We have

$$E(X)=(1+2+3+4+5+6)\cdot\frac{1}{6}=\frac{21}{6}=\frac{7}{2}$$

and

$$E(X^2)=(1^2+2^2+3^2+4^2+5^2+6^2)\cdot\frac{1}{6}=\frac{91}{6}.$$

Thus,

$$\text{Var}(X)=\frac{91}{6}-\frac{49}{4}=\frac{35}{12}$$

The standard deviation is

$$\sigma_X = \sqrt{\frac{35}{12}} \approx 1.7078 \blacksquare$$

Next, we will show that $g(c) = E[(X-c)^2]$ is minimum when $c = E(X)$.

Theorem 17.2
Let c be a constant and let X be a random variable with mean $E(X)$ and variance $\text{Var}(X) < \infty$. Then
(a) $g(c) = E[(X-c)^2] = \text{Var}(X) + (c - E(X))^2$.
(b) $g(c)$ is minimized at $c = E(X)$.

Proof.
(a) We have

$$\begin{aligned}
E[(X-c)^2] =& E[((X - E(X)) - (c - E(X)))^2] \\
=& E[(X - E(X))^2] - 2(c - E(X))E(X - E(X)) + (c - E(X))^2 \\
=& \text{Var}(X) + (c - E(X))^2
\end{aligned}$$

(b) Note that $g'(c) = 0$ when $c = E(X)$ and $g''(E(X)) = 2 > 0$ so that g has a global minimum at $c = E(X)$ $\blacksquare$

Practice Problems

Problem 17.1 ‡
A probability distribution of the claim sizes for an auto insurance policy is given in the table below:

Claim size	Probability
20	0.15
30	0.10
40	0.05
50	0.20
60	0.10
70	0.10
80	0.30

What percentage of the claims are within one standard deviation of the mean claim size?

Problem 17.2 ‡
The annual cost of maintaining and repairing a car averages 200 with a variance of 260. what will be the variance of the annual cost of maintaining and repairing a car if 20% tax is introduced on all items associated with the maintenance and repair of cars?

Problem 17.3
A discrete random variable, X, has probability mass function

$$p(x) = c(x-3)^2, \quad x = -2, -1, 0, 1, 2.$$

(a) Find the value of the constant c.
(b) Find the mean and variance of X.

Problem 17.4
An urn contains 10 marbles in which 3 are black. Four of the marbles are selected at random and are tested for the black color. Define the random variable X to be the number of the selected marbles that are not black.
(a) Find the probability mass function of X.
(b) Find the variance of X.

Problem 17.5
Suppose that X is a discrete random variable with probability mass function

$$p(x) = cx^2, \quad x = 1, 2, 3, 4.$$

(a) Find the value of c.
(b) Find $E(X)$ and $E(X(X-1))$.
(c) Find $\mathrm{Var}(X)$.

Problem 17.6
Suppose X is a random variable with $E(X) = 4$ and $\mathrm{Var}(X) = 9$. Let $Y = 4X + 5$. Compute $E(Y)$ and $\mathrm{Var}(Y)$.

Problem 17.7
A box contains 3 red and 4 blue marbles. Two marbles are randomly selected without replacement. If they are the same color then you win \$2. If they are of different colors then you lose \$ 1. Let X denote the amount you win.
(a) Find the probability mass function of X.
(b) Compute $E(X)$ and $E(X^2)$.
(c) Find $\mathrm{Var}(X)$.

Problem 17.8
Let X be a discrete random variable with probability mass function is given by

x	-4	1	4
$p(x)$	0.3	0.4	0.3

Find the variance and the standard deviation of X.

Problem 17.9
Let X be a random variable with probability distribution $p(0) = 1-p, p(1) = p$, and 0 otherwise, where $0 < p < 1$. Find $E(X)$ and $Var(X)$.

Commonly Used Discrete Random Variables

In this chapter, we consider the discrete random variables listed in the exam's syllabus: binomial, negative binomial, geometric, hypergeometric, and Poisson.

18 Bernoulli Trials and Binomial Distributions

A **Bernoulli trial**[3] is an experiment with exactly two outcomes: **Success** and **failure.** The probability of a success is denoted by p and that of a failure by q. Moreover, p and q are related by the formula

$$p + q = 1.$$

Example 18.1
Consider the experiment of rolling a fair die where a success is the face that comes up shows a number divisible by 2. Find p and q.

Solution.
The numbers on the die that are divisible by 2 are 2,4, and 6. Thus, $p = \frac{3}{6} = \frac{1}{2}$ and $q = 1 - \frac{1}{2} = \frac{1}{2}$ ■
A **Bernoulli experiment** is a sequence of independent[4] Bernoulli trials.
Let X represent the number of successes that occur in n indepednent Bernoulli trials. Then X is said to be a **binomial random variable** with parameters

[3]The prefix bi in binomial experiment refers to the fact that there are two possible outcomes (e.g., head or tail, true or false, working or defective) to each trial.
[4]That is what happens to one trial does not affect the probability of a success in any other trial.

(n, p). If $n = 1$ then X is said to be a **Bernoulli random variable**.
The central question of a binomial experiment is to find the probability of r successes out of n trials. In the next paragraph we'll see how to compute such a probability. Now, anytime we make selections from a population **without replacement,** we do not have independent trials. For example, selecting a ball from a box that contain balls of two different colors.

Example 18.2
We roll a fair die 5 times. A success is when the face that comes up shows a prime number. We are interested in the probability of obtaining three prime numbers. What are p, q, n, and r?

Solutions.
This is a binomial experiment with 5 trials. The prime numbers on the die are $2, 3, 5$ so that $p = q = \frac{1}{2}$. Also, we have $n = 5$ and $r = 3$ ■

Binomial Distribution Function
As mentioned above, the central problem of a binomial experiment is to find the probability of r successes out of n independent trials.

Recall from Section 5 the formula for finding the number of combinations of n distinct objects taken r at a time

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

Now, the probability of r successes in a sequence of n independent trials is given by $p^r q^{n-r}$. Since the binomial coefficient $_nC_r$ counts all the number of outcomes that have r successes and $n - r$ failures, the probability of having r successes in any order is given by the **binomial mass function**

$$p(r) = \Pr(X = r) = {}_nC_r p^r q^{n-r}.$$

Note that by letting $a = p$ and $b = 1 - p$ in the binomial formula we find

$$\sum_{k=0}^{n} p(k) = \sum_{k=0}^{n} {}_nC_k p^k (1-p)^{n-k} = (p + 1 - p)^n = 1.$$

The histogram of the binomial distribution is given in Figure 18.1.

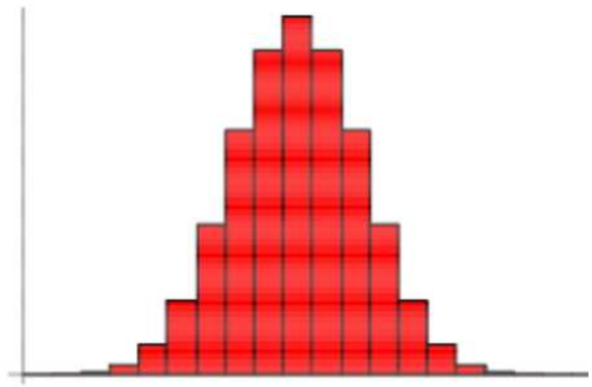


Figure 18.1

The **cumulative distribution function** is given by

$$F(x) = \Pr(X \leq x) = \begin{cases} 0, & x < 0 \\ \sum_{k=0}^{\lfloor x \rfloor} {}_nC_k p^k (1-p)^{n-k}, & 0 \leq x \leq n \\ 1, & x > n \end{cases}$$

where $\lfloor k \rfloor$ is the **floor function**[5].

Example 18.3
Suppose that in a box of 100 computer chips, the probability of a chip to be defective is 3%. Inspection process for defective chips consists of selecting with replacement 5 randomly chosen chips in the box and to send the box for shipment if none of the five chips is defective. Write down the random variable, the corresponding probability distribution and then determine the probability that the box described here will be allowed to be shipped.

Solution.
Let X be the number of defective chips in the box. Then X is a binomial random variable with probability distribution

$$\Pr(X = x) = {}_5C_x (0.03)^x (0.97)^{5-x}, \; x = 0, 1, 2, 3, 4, 5.$$

[5] $\lfloor x \rfloor$ = the largest integer less than or equal to x.

Now,

$$\Pr(\text{sheet goes into circulation}) = \Pr(X = 0) = (0.97)^5 = 0.859 \blacksquare$$

Example 18.4
Suppose that 40% of the voters in a city are in favor of a ban of smoking in public buildings. Suppose 5 voters are to be randomly sampled. Find the probability that
(a) 2 favor the ban.
(b) less than 4 favor the ban.
(c) at least 1 favor the ban.

Solution.
(a) $\Pr(X = 2) = {}_5C_2(0.4)^2(0.6)^3 \approx 0.3456.$
(b) $\Pr(X < 4) = p(0)+p(1)+p(2)+p(3) = {}_5C_0(0.4)^0(0.6)^5+{}_5C_1(0.4)^1(0.6)^4+ {}_5C_2(0.4)^2(0.6)^3 + {}_5C_3(0.4)^3(0.6)^2 \approx 0.913.$
(c) $\Pr(X \geq 1) = 1 - \Pr(X < 1) = 1 - {}_5C_0(0.4)^0(0.6)^5 \approx 0.922 \blacksquare$

Example 18.5
A student takes a test consisting of 10 true-false questions.
(a) What is the probability that the student answers at least six questions correctly?
(b) What is the probability that the student answers at most two questions correctly?

Solution.
(a) Let X be the number of correct responses. Then X is a binomial random variable with parameters $n = 10$ and $p = \frac{1}{2}$. So, the desired probability is

$$\begin{aligned}\Pr(X \geq 6) =&\Pr(X = 6) + \Pr(X = 7) + \Pr(X = 8) + \Pr(X = 9) + \Pr(X = 10)\\ =&\sum_{x=6}^{10} {}_{10}C_x(0.5)^x(0.5)^{10-x} \approx 0.3769.\end{aligned}$$

(b) We have

$$\Pr(X \leq 2) = \sum_{x=0}^{2} {}_{10}C_x(0.5)^x(0.5)^{10-x} \approx 0.0547 \blacksquare$$

Example 18.6
A study shows that 30 percent people aged 50-60 in a certain town have high blood pressure. What is the probability that in a sample of fourteen individuals aged between 50 and 60 tested for high blood pressure, more than six will have high blood pressure?

Solution.
Let X be the number of people in the town aged 50-60 with high blood pressure. Then X is a binomial random variable with $n = 14$ and $p = 0.3$. Thus,

$$\begin{aligned}\Pr(X > 6) =&1 - \Pr(X \leq 6)\\ =&1 - \sum_{i=0}^{6} {}_{14}C_i (0.3)^i (0.7)^{14-i}\\ \approx&0.0933 \blacksquare\end{aligned}$$

Practice Problems

Problem 18.1
Mark is a car salesman with a 10% chance of persuading a randomly selected customer to buy a car. Out of 8 customers that were serviced by Mark, what is the probability that exactly one agreed to buy a car?

Problem 18.2
The probability of a newly born child to get a genetic disease is 0.25. If a couple carry the disease and wish to have four children then what is the probability that 2 of the children will get the disease?

Problem 18.3
A skyscraper has three elevators. Each elevator has a 50% chance of being down, independently of the others. Let X be the number of elevators which are down at a particular time. Find the probability mass function (pmf) of X.

Problem 18.4 ‡
A hospital receives 1/5 of its flu vaccine shipments from Company X and the remainder of its shipments from other companies. Each shipment contains a very large number of vaccine vials.
For Company Xs shipments, 10% of the vials are ineffective. For every other company, 2% of the vials are ineffective. The hospital tests 30 randomly selected vials from a shipment and finds that one vial is ineffective.
What is the probability that this shipment came from Company X?

Problem 18.5 ‡
A company establishes a fund of 120 from which it wants to pay an amount, C, to any of its 20 employees who achieve a high performance level during the coming year. Each employee has a 2% chance of achieving a high performance level during the coming year, independent of any other employee.
Determine the maximum value of C for which the probability is less than 1% that the fund will be inadequate to cover all payments for high performance.

Problem 18.6 ‡
A company prices its hurricane insurance using the following assumptions:

(i) In any calendar year, there can be at most one hurricane.
(ii) In any calendar year, the probability of a hurricane is 0.05 .
(iii) The number of hurricanes in any calendar year is independent of the number of hurricanes in any other calendar year.

Using the company's assumptions, calculate the probability that there are fewer than 3 hurricanes in a 20-year period.

Problem 18.7
The probability of winning a game is $\frac{1}{300}$. If you play this game 200 times, what is the probability that you win at least twice?

Problem 18.8
Suppose a local bus service accepted 12 reservations for a commuter bus with 10 seats. Seven of the ten reservations went to regular commuters who will show up for sure. The other 5 passengers will show up with a 50% chance, independently of each other.
(a) Find the probability that the bus will be overbooked.
(b) Find the probability that there will be empty seats.

Problem 18.9
Suppose that 3% of flashlight batteries produced by a certain machine are defective. The batteries are put into packages of 20 batteries for distribution to retailers.
What is the probability that a randomly selected package of batteries will contain at least 2 defective batteries?

Problem 18.10
The probability of late arrival of flight 701 in any day is 0.20 and is independent of the late arrival in any other day. The flight can be late only once per day. Calculate the probability that the flight is late two or more times in ten days.

Problem 18.11
Ashley finds that she beats Carla in tennis 70% of the time. The two play 3 times in a particular month. Assuming independence of outcomes, what is the probability Ashley wins at least 2 of the 3 matches?

Problem 18.12
The probability of a computer chip to be defective is 0.05. Consider a package of 6 computer chips.
(a) What is the probability one chip will be defective?
(b) What is the probability at least one chip will be defective?
(c) What is the probability that more than one chip will be defective, given at least one is defective?

Problem 18.13
In a promotion, a popcorn company inserts a coupon for a free Red Box movie in 10% of boxes produced. Suppose that we buy 10 boxes of popcorn, what is the probability that we get at least 2 coupons?

19 The Expected Value and Variance of the Binomial Distribution

In this section, we find the expected value and the variance of a binomial random variable X with parameters (n, p).

The expected value is found as follows.

$$\begin{aligned}
E(X) =& \sum_{k=1}^{n} k\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)!(n-k)!}p^{k-1}(1-p)^{n-k}\\
=& np\sum_{j=0}^{n-1}\frac{(n-1)!}{j!(n-1-j)!}p^j(1-p)^{n-1-j} = np(p+1-p)^{n-1} = np
\end{aligned}$$

where we used the binomial theorem and the substitution $j = k-1$. Also, we have

$$\begin{aligned}
E(X(X-1)) =& \sum_{k=0}^{n} k(k-1)\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}\\
=& n(n-1)p^2\sum_{k=2}^{n}\frac{(n-2)!}{(k-2)!(n-k)!}p^{k-2}(1-p)^{n-k}\\
=& n(n-1)p^2\sum_{j=0}^{n-2}\frac{(n-2)!}{j!(n-2-j)!}p^j(1-p)^{n-2-j}\\
=& n(n-1)p^2(p+1-p)^{n-2} = n(n-1)p^2
\end{aligned}$$

This implies $E(X^2) = E(X(X-1)) + E(X) = n(n-1)p^2 + np$. The variance of X is then

$$\text{Var}(X) = E(X^2) - (E(X))^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)$$

Example 19.1
The probability of a student passing an exam is 0.2. Ten students took the exam.
(a) What is the probability that at least two students passed the exam?
(b) What is the expected number of students who passed the exam?
(c) How many students must take the exam to make the probability at least 0.99 that a student will pass the exam?

Solution.
Let X be the number of students who passed the exam. Then, X has a binomial distribution with $n = 10$ and $p = 0.2$.
(a) The event that at least two students passed the exam is $\{X \geq 2\}$. So,

$$\begin{aligned}\Pr(X \geq 2) =& 1 - \Pr(X < 2) = 1 - p(0) - p(1)\\ =& 1 - {}_{10}C_0(0.2)^0(0.8)^{10} - {}_{10}C_1(0.2)^1(0.8)^9\\ \approx& 0.6242\end{aligned}$$

(b) $E(X) = np = 10 \cdot (0.2) = 2$.
(c) Suppose that n students are needed to make the probability at least 0.99 that a student will pass the exam. Let A denote the event that a student pass the exam. Then, A^c means that all the students fail the exam. We have,

$$\Pr(A) = 1 - \Pr(A^c) = 1 - (0.8)^n \geq 0.99$$

Solving the inequality, we find that $n \geq \frac{\ln(0.01)}{\ln(0.8)} \approx 20.6$. So, the required number of students is 21 ■

Example 19.2
Let X be a binomial random variable with parameters $(12, 0.5)$. Find the variance and the standard deviation of X.

Solution.
We have $n = 12$ and $p = 0.5$. Thus, $\text{Var}(X) = np(1-p) = 6(1-0.5) = 3$. The standard deviation is $\sigma_X = \sqrt{3}$ ■

Example 19.3
A multiple choice exam consists of 25 questions each with five choices with once choice is correct. Randomly select an answer for each question. Let X be the random variable representing the total number of correctly answered questions.
(a) What is the probability that you get exactly 16, or 17, or 18 of the questions correct?
(b) What is the probability that you get at least one of the questions correct.
(c) Find the expected value of the number of correct answers.

Solution.
(a) Let X be the number of correct answers. We have

$$\begin{aligned}\Pr(X = 16 \text{ or } X = 17 \text{ or } X = 18) =& {}_{25}C_{16}(0.2)^{16}(0.8)^9 + {}_{25}C_{17}(0.2)^{17}(0.8)^8\\ &+ {}_{25}C_{18}(0.2)^{18}(0.8)^7 = 2.06 \times 10^{-6}.\end{aligned}$$

(b) $\Pr(X \geq 1) = 1 - \Pr(X = 0) = 1 - {}_{25}C_0(0.8)^{25} = 0.9962.$
(c) We have $E(X) = 25(0.2) = 5$ ■

A useful fact about the binomial distribution is a recursion for calculating the probability mass function.

Theorem 19.1
Let X be a binomial random variable with parameters (n, p). Then for $k = 1, 2, 3, \cdots, n$

$$p(k) = \frac{p}{1-p}\frac{n-k+1}{k}p(k-1)$$

Proof.
We have

$$\begin{aligned}\frac{p(k)}{p(k-1)} =& \frac{{}_nC_k p^k(1-p)^{n-k}}{{}_nC_{k-1}p^{k-1}(1-p)^{n-k+1}}\\ =& \frac{\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}}{\frac{n!}{(k-1)!(n-k+1)!}p^{k-1}(1-p)^{n-k+1}}\\ =& \frac{(n-k+1)p}{k(1-p)} = \frac{p}{1-p}\frac{n-k+1}{k}\end{aligned} ■$$

The following theorem details how the binomial pmf first increases and then decreases.

Theorem 19.2
Let X be a binomial random variable with parameters (n, p). As k goes from 0 to n, $p(k)$ first increases monotonically and then decreases monotonically reaching its largest value when k is the largest integer such that $k \leq (n+1)p$.

Proof.
From the previous theorem we have

$$\frac{p(k)}{p(k-1)} = \frac{p}{1-p}\frac{n-k+1}{k} = 1 + \frac{(n+1)p - k}{k(1-p)}.$$

Accordingly, $p(k) > p(k-1)$ when $k < (n+1)p$ and $p(k) < p(k-1)$ when $k > (n+1)p$. Now, if $(n+1)p = m$ is an integer then $p(m) = p(m-1)$. If not, then by letting $k = [(n+1)p]$ = greatest integer less than or equal to

$(n+1)p$ we find that p reaches its maximum at k ■

We illustrate the previous theorem by a histogram.

Binomial Random Variable Histogram
The histogram of a binomial random variable is constructed by putting the r values on the horizontal axis and $p(r)$ values on the vertical axis. The width of the bar is 1 and its height is $p(r)$. The bars are centered at the r values.

Example 19.4
Construct the binomial distribution for the total number of heads in four flips of a balanced coin. Make a histogram.

Solution.
The binomial distribution is given by the following table

r	0	1	2	3	4
$p(r)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

The corresponding histogram is shown in Figure 19.1 ■

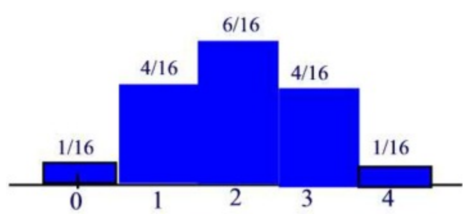


Figure 19.1

Practice Problems

Problem 19.1
If X is the number of "6"'s that turn up when 72 ordinary dice are independently thrown, find the expected value of X^2.

Problem 19.2 ‡
A tour operator has a bus that can accommodate 20 tourists. The operator knows that tourists may not show up, so he sells 21 tickets. The probability that an individual tourist will not show up is 0.02, independent of all other tourists.
Each ticket costs 50, and is non-refundable if a tourist fails to show up. If a tourist shows up and a seat is not available, the tour operator has to pay 100 (ticket cost + 50 penalty) to the tourist.
What is the expected revenue of the tour operator?

Problem 19.3
Let Y be a binomial random variable with parameters $(n, 0.2)$. Define the random variable

$$S = 100 + 50Y - 10Y^2.$$

Give the expected value of S when $n = 1$, 2, and 3.

Problem 19.4
A recent study shows that the probability of a marriage will end in a divorce within 10 years is 0.4. Letting a divorce be a success, find the mean and the standard deviation for the binomial distribution X involving 1000 marriages.

Problem 19.5
The probability of a person contracting the flu on exposure is 0.4. Let a success be a person contracting the flu. Consider the binomial distribution for a group of 5 people that has been exposed.
(a) Find the probability mass function.
(b) Compute $p(x)$ for $x = 0, 1, 2, 3, 4, 5$.
(c) Draw a histogram for the distribution.
(d) Find the mean and the standard deviation.

Problem 19.6
A fair die is rolled twice. A success is when the face that comes up shows 3

or 6. (a) Write the function defining the distribution.
(b) Construct a table for the distribution.
(c) Construct a histogram for the distribution.
(d) Find the mean and the standard deviation for the distribution.

20 Poisson Random Variable

A random variable X is said to be a **Poisson** random variable with parameter $\lambda > 0$ if its probability mass function has the form

$$p(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \cdots$$

where λ indicates the average number of successes per unit time or space. Note that $p(k) \geq 0$ and

$$\sum_{k=0}^{\infty} p(k) = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1.$$

The Poisson random variable is most commonly used to model the number of random occurrences of some phenomenon in a specified unit of space or time. For example, the number of phone calls received by a telephone operator in a 10-minute period or the number of typos per page made by a secretary.

Example 20.1
The number of car accidents on a certain section of highway I40 averages 2.1 per day. Assuming that the number of accidents in a given day follows a Poisson distribution, what is the probability that 4 accidents will occur on a given day?

Solution.
The probability that 4 accidents will occur on a given day is given by

$$\Pr(X = 4) = e^{-2.1}\frac{(2.1)^4}{4!} \approx 0.0992 \blacksquare$$

Example 20.2
The number of people entering a movie theater averages one every two minutes. Assuming that a Poisson distribution is appropriate.
(a) What is the probability that no people enter between 12:00 and 12:05?
(b) Find the probability that at least 4 people enter during [12:00,12:05].

Solution.
(a) Let X be the number of people that enter between 12:00 and 12:05. We model X as a Poisson random variable with parameter λ, the average number

of people that arrive in the 5-minute interval. But, if 1 person arrives every 2 minutes, on average (so 1/2 a person per minute), then in 5 minutes an average of 2.5 people will arrive. Thus, $\lambda = 2.5$. Now,

$$\Pr(X=0) = e^{-2.5}\frac{2.5^0}{0!} = e^{-2.5}.$$

(b)

$$\begin{aligned}\Pr(X \geq 4) =& 1 - \Pr(X \leq 3)\\ =& 1 - \Pr(X=0) - \Pr(X=1) - \Pr(X=2) - \Pr(X=3)\\ =& 1 - e^{-2.5}\frac{2.5^0}{0!} - e^{-2.5}\frac{2.5^1}{1!} - e^{-2.5}\frac{2.5^2}{2!} - e^{-2.5}\frac{2.5^3}{3!} \blacksquare\end{aligned}$$

Example 20.3
The number of weekly life insurance sold by an insurance agent averages 3 per week. Assuming that this number follows a Poisson distribution, calculate the probability that in a given week the agent will sell
(a) some policies
(b) 2 or more policies but less than 5 policies.
(c) Assuming that there are 5 working days per week, what is the probability that in a given day the agent will sell one policy?

Solution.
(a) Let X be the number of policies sold in a week. Then

$$\Pr(X \geq 1) = 1 - \Pr(X=0) = 1 - \frac{e^{-3}3^0}{0!} \approx 0.95021$$

(b) We have

$$\begin{aligned}\Pr(2 \leq X < 5) =& \Pr(X=2) + \Pr(X=3) + \Pr(X=4)\\ =& \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!} + \frac{e^{-3}3^4}{4!} \approx 0.61611\end{aligned}$$

(c) Let X be the number of policies sold per day. Then $\lambda = \frac{3}{5} = 0.6$. Thus,

$$\Pr(X=1) = \frac{e^{-0.6}(0.6)}{1!} \approx 0.32929 \blacksquare$$

Example 20.4
The number X of people in a certain town who get the influenza follows a Poisson distribution. The proportion of people who did not get the flu is 0.01. Find the probability mass function of X.

Solution.
The pmf of X is

$$\Pr(X=k)=\lambda^k\frac{e^{-\lambda}}{k!},\quad k=0,1,2,\cdots$$

Since $P(X=0)=0.01$ we can write $e^{-\lambda}=0.01$. Thus, $\lambda=4.605$ and the exact Poisson distribution is

$$\Pr(X=k)=(4.605)^k\frac{e^{-4.605}}{k!},\quad k=0,1,2,\cdots \blacksquare$$

If X has a Poissom distribution, its expected value is found as follows.

$$\begin{aligned}E(X)=&\sum_{k=1}^{\infty}k\frac{e^{-\lambda}\lambda^k}{k!}=\lambda\sum_{k=1}^{\infty}\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\\=&\lambda\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}=\lambda e^{-\lambda}e^{\lambda}=\lambda\end{aligned}$$

To find the variance, we first compute $E(X^2)$. From

$$\begin{aligned}E(X(X-1))=&\sum_{k=2}^{\infty}k(k-1)\frac{e^{-\lambda}\lambda^k}{k!}=\lambda^2\sum_{k=2}^{\infty}\frac{e^{-\lambda}\lambda^{k-2}}{(k-2)!}\\=&\lambda^2\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}=\lambda^2e^{-\lambda}e^{\lambda}=\lambda^2\end{aligned}$$

we find $E(X^2)=E(X(X-1))+E(X)=\lambda^2+\lambda$. Thus, $\mathrm{Var}(X)=E(X^2)-(E(X))^2=\lambda$.

Example 20.5
Misprints in a book averages one misprint per 10 pages. Suppose that the number of misprints is a random variable having Poisson distribution. Let X denote the number of misprints in a stack of 50 pages. Find the mean and the standard deviation of X.

Solution.
Since the book has 1 misprint per 10 pages, the number of misprints in a stack of 50 pages is 5. Thus, X is a Poisson random variable with parameter $\lambda = 5$. Hence, $E(X) = \lambda = 5$ and $\sigma_X = \sqrt{5}$ ■

Example 20.6 ‡
Let X represent the number of customers arriving during the morning hours and let Y represent the number of customers arriving during the afternoon hours at a diner. You are given:
i) X and Y are Poisson distributed.
ii) The first moment of X is less than the first moment of Y by 8.
iii) The second moment of X is 60% of the second moment of Y.
Calculate the variance of Y.

Solution.
We have

$$E(X) = E(Y) - 8 \Longrightarrow E(X)^2 = E(Y)^2 - 16E(Y) + 64.$$

But

$$E(X)^2 = \mathrm{Var}(X) + E(X) = E(X) + E(X)^2 = 0.6E(Y^2) = 0.6(E(Y) + E(Y)^2).$$

Hence,

$$0.6(E(Y) + E(Y)^2) = E(Y)^2 - 16E(Y) + 64$$

yields the quadratic equation

$$0.4E(Y)^2 - 15.5E(Y) + 56 = 0$$

whose roots are $E(Y) = 4$ and $E(Y) = 35$. The value $E(Y) = 4$ yields $E(X) = 4 - 8 = -4$ which is impossible since the expected value of a Poisson random variable is always positive. Hence, $E(Y) = 35$ ■

Practice Problems

Problem 20.1
The number of accidents on a certain section of a highway averages 4 per day. Assuming that this number follows a Poisson distribution, what is the probability of no car accident in one day? What is the probability of 1 car accident in two days?

Problem 20.2
A phone operator receives calls on average of 2 calls per minute. What is the probability of receiving 10 calls in 5 minutes?

Problem 20.3
In the first draft of a book on probability theory, there are an average of 15 spelling errors per page. Suppose that the number of errors per page follows a Poisson distribution. What is the probability of having no errors on a page?

Problem 20.4
Suppose that the number of people admitted to an emergency room each day is a Poisson random variable with parameter $\lambda = 3$.
(a) Find the probability that 3 or more people admitted to the emergency room today.
(b) Find the probability that no people were admitted to the emergency room today.

Problem 20.5
At a reception event guests arrive at an average of 2 per minute. Find the probability that
(a) at most 4 will arrive at any given minute
(b) at least 3 will arrive during an interval of 2 minutes
(c) 5 will arrive in an interval of 3 minutes.

Problem 20.6
Suppose that the number of car accidents on a certain section of a highway can be modeled by a random variable having Poisson distribution with standard deviation $\sigma = 2$. What is the probability that there are at least three accidents?

Problem 20.7
A Geiger counter is monitoring the leakage of alpha particles from a container of radioactive material. Over a long period of time, an average of 50 particles per minute is measured. Assume the arrival of particles at the counter is modeled by a Poisson distribution.
(a) Compute the probability that at least one particle arrives in a particular one second period.
(b) Compute the probability that at least two particles arrive in a particular two second period.

Problem 20.8 ‡
An actuary has discovered that policyholders are three times as likely to file two claims as to file four claims.
If the number of claims filed has a Poisson distribution, what is the variance of the number of claims filed?

Problem 20.9 ‡
A company buys a policy to insure its revenue in the event of major snowstorms that shut down business. The policy pays nothing for the first such snowstorm of the year and $10,000 for each one thereafter, until the end of the year. The number of major snowstorms per year that shut down business is assumed to have a Poisson distribution with mean 1.5 .
What is the expected amount paid to the company under this policy during a one-year period?

Problem 20.10 ‡
A baseball team has scheduled its opening game for April 1. If it rains on April 1, the game is postponed and will be played on the next day that it does not rain. The team purchases insurance against rain. The policy will pay 1000 for each day, up to 2 days, that the opening game is postponed.
The insurance company determines that the number of consecutive days of rain beginning on April 1 is a Poisson random variable with mean 0.6 .
What is the standard deviation of the amount the insurance company will have to pay?

Problem 20.11
The average number of trains arriving on any one day at a train station in a certain city is known to be 12. What is the probability that on a given day fewer than nine trains will arrive at this station?

Problem 20.12
In the inspection sheet metals produced by a machine, five defects per 10 square feet were spotted, on average. If we assume a Poisson distribution, what is the probability that a 15-square feet sheet of the metal will have at least six defects?

Problem 20.13
Let X be a Poisson random variable with mean λ. If $P(X = 1|X \leq 1) = 0.8$, what is the value of λ?

Problem 20.14
The number of trucks arriving at a truck depot on a given day has a Poisson distribution with a mean of 2.5 per day.
(a) What is the probability a day goes by with no more than one truck arriving?
(b) Give the mean and standard deviation of the number of trucks arriving in an 8-day period.

21 Poisson Approximation to the Binomial Distribution

In this section, we show that a binomial random variable with parameters n and p such that n is large and p is small can be approximated by a Poisson distribution.

Theorem 21.1
Let X be a binomial random variable with parameters n and p. If $n \to \infty$ and $p \to 0$ so that $np = \lambda = E(X)$ remains constant then X can be approximated by a Poisson distribution with parameter λ.

Proof.
First notice that for small $p << 1$ we can write

$$\begin{aligned}(1-p)^n =& e^{n\ln(1-p)}\\ =& e^{n(-p-\frac{p^2}{2}-\cdots)}\\ \approx& e^{-np}\\ =& e^{-\lambda}\end{aligned}$$

where we have used the Taylor series expansion

$$\ln(1-x) = \sum_{n=1}^{\infty}(-1)^{2n+1}\frac{x^n}{n}, \ |x| \le 1.$$

We prove that

$$\Pr(X=k) \approx e^{-\lambda}\frac{\lambda^k}{k!}.$$

This is true for $k=0$ since $\Pr(X=0) = (1-p)^n \approx e^{-\lambda}$. Suppose $k > 0$. Using the fact that

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x$$

we find

$$\begin{aligned}
\Pr(X=k) =&{}_nC_k\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}\\
=&\frac{n(n-1)\cdots(n-k+1)}{n^k}\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}\\
=&\frac{n}{n}\cdot\frac{n-1}{n}\cdots\frac{n-k+1}{n}\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}\\
\to&1\cdot\frac{\lambda^k}{k!}\cdot e^{-\lambda}\cdot 1
\end{aligned}$$

as $n\to\infty$. Note that for $0\le j\le k-1$ we have $\frac{n-j}{n}=1-\frac{j}{n}\to 1$ as $n\to\infty$ ■

In general, Poisson distribution will provide a good approximation to binomial probabilities when $n\ge 20$ and $p\le 0.05$. When $n\ge 100$ and $p\le 0.01$, the approximation will generally be excellent.

Example 21.1
In a group of 100 individuals, let X be the random variable representing the total number of people in the group with a birthday on Thanksgiving day. Then X is a binomial random variable with parameters $n=100$ and $p=\frac{1}{365}$. What is the probability at least one person in the group has a birthday on Thanksgiving day?

Solution.
We have

$$\Pr(X\ge 1)=1-\Pr(X=0)=1-{}_{100}C_0\left(\frac{1}{365}\right)^0\left(\frac{364}{365}\right)^{100}\approx 0.2399.$$

Using the Poisson approximation, with $\lambda=100\times\frac{1}{365}=\frac{100}{365}=\frac{20}{73}$ we find

$$\Pr(X\ge 1)=1-\Pr(X=0)=1-\frac{(20/73)^0}{0!}e^{\left(-\frac{20}{73}\right)}\approx 0.2396 ■$$

Example 21.2
Consider the experiment of rolling two fair dice 6 times. Let X denote the number of times a double 4 appears. For $k=0,1,2$ compare $\Pr(X=k)$ found using the binomial distribution with the one found using Poisson approximation.

Solution.
We are given that $n = 6$ and $p = \frac{1}{36}$ so that $\lambda = np = \frac{1}{6}$. Let $B(X = k)$ be the probability using binomial distribution and $P(X = k)$ be the probability using Poisson distribution. Then

$$\begin{aligned}
B(X = 0) =& \left(1 - \frac{1}{36}\right)^6 \approx 0.8445\\
P(X = 0) =& e^{-\frac{1}{6}} \approx 0.8465\\
B(X = 1) =& {}_6C_1 \frac{1}{36}\left(1 - \frac{1}{36}\right)^5 \approx 0.1448\\
P(X = 1) =& e^{-\frac{1}{6}}\left(\frac{1}{6}\right) \approx 0.1411\\
B(X = 2) =& {}_6C_2 \left(\frac{1}{36}\right)^2\left(1 - \frac{1}{36}\right)^4 \approx 0.0103\\
P(X = 2) =& e^{-\frac{1}{6}}\frac{\left(\frac{1}{6}\right)^2}{2!} \approx 0.0118 \blacksquare
\end{aligned}$$

Example 21.3
Consider a contest where a participant fires at a small can placed on the top of box. Each time the can is hit, it is replaced by another can. Suppose that the probability of a paticipant hiting the can is $\frac{1}{32}$. Assume that the participant shoots 96 times, and that all shoots are independent.
(a) Find the probability mass function of the number of shoots that hit a can.
(b) Give an approximation for the probability of the participant hitting no more than one can.

Solution.
(a) Let X denote the number of shoots that hit a can. Then X is binomially distributed:

$$\Pr(X = k) = {}_nC_k p^k(1 - p)^{n-k}, \quad n = 96,\ p = \frac{1}{32}.$$

(b) Since n is large, and p small, we can use the Poisson approximation, with parameter $\lambda = np = 3$. Thus,

$$\Pr(X \le 1) = \Pr(X = 0) + \Pr(X = 1) \approx e^{-\lambda} + \lambda e^{-\lambda} = 4e^{-3} \approx 0.199 \blacksquare$$

We conclude this section by establishing a recursion formula for computing $p(k) = \Pr(X = k)$.

Theorem 21.2
If X is a Poisson random variable with parameter λ, then

$$p(k+1) = \frac{\lambda}{k+1}p(k).$$

Proof.
We have

$$\begin{aligned}
\frac{p(k+1)}{p(k)} =& \frac{e^{-\lambda}\frac{\lambda^{k+1}}{(k+1)!}}{e^{-\lambda}\frac{\lambda^k}{k!}} \\
=& \frac{\lambda}{k+1} \blacksquare
\end{aligned}$$

Practice Problems

Problem 21.1
Let X be a binomial distribution with parameters $n = 200$ and $p = 0.02$. We want to calculate $\Pr(X \geq 2)$. Explain why a Poisson distribution can be expected to give a good approximation of $\Pr(X \geq 2)$ and then find the value of this approximation.

Problem 21.2
In a TV plant, the probability of manufacturing a defective TV is 0.03. Using Poisson approximation, find the probability of obtaining exactly one defective TV set out of a group of 20.

Problem 21.3
Suppose that 1 out of 400 tires are devective. Let X denote the number of defective tires in a group of 200 tires. What is the probability that at least three of them are defective?

Problem 21.4
1000 cancer patients are receving a clinical trial drug for cancer. Side effects are being studied. The probability that a patient experiences side effects to the drug is found to be 0.001. Find the probability that none of the patients administered the trial drug experienced any side effect.

Problem 21.5
From a group of 120 engineering students, 3% are not in favor of studying differential equations. Use the Poisson approximation to estimate the probability that
(a) exactly 2 students are not in favor of studying differential equations;
(b) at least two students are not in favor of studying differential equations.

22 Geometric Random Variable

A **geometric** random variable with parameter $p, 0 < p < 1$ has a probability mass function

$$p(n) = \Pr(X = n) = p(1-p)^{n-1}, \quad n = 1, 2, \cdots .$$

Note that $p(n) \geq 0$ and

$$\sum_{n=1}^{\infty} p(1-p)^{n-1} = \frac{p}{1-(1-p)} = 1.$$

A geometric random variable models the number of successive independent Bernoulli trials that must be performed to obtain the first "success". For example, the number of flips of a fair coin until the first head appears follows a geometric distribution.

Example 22.1
Consider the experiment of rolling a pair of fair dice.
(a) What is the probability of getting a sum of 11?
(b) If you roll the dice repeatedly, what is the probability that the first 11 occurs on the 8^{th} roll?

Solution.
(a) A sum of 11 accurs when the pair of dice show either $(5,6)$ or $(6,5)$ so that the required probability is $\frac{2}{36} = \frac{1}{18}$.
(b) Let X be the number of rolls on which the first 11 occurs. Then X is a geometric random variable with parameter $p = \frac{1}{18}$. Thus,

$$\Pr(X = 8) = \left(\frac{1}{18}\right)\left(1 - \frac{1}{18}\right)^7 = 0.0372 \blacksquare$$

To find the expected value and variance of a geometric random variable we proceed as follows. First we recall from calculus the geometric series

$$f(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \; |x| < 1$$

Differentiating $f(x)$ twice we find

$$f'(x) = \sum_{n=1}^{\infty} nx^{n-1} = (1-x)^{-2} \quad \text{and} \quad f''(x) = \sum_{n=1}^{\infty} n(n-1)x^{n-2} = 2(1-x)^{-3}.$$

Evaluating $f'(x)$ and $f''(x)$ at $x = 1-p$, we find

$$f'(1-p) = \sum_{n=1}^{\infty} n(1-p)^{n-1} = p^{-2}$$

and

$$f''(1-p) = \sum_{n=1}^{\infty} n(n-1)(1-p)^{n-2} = 2p^{-3}.$$

We next apply these equalities in finding $E(X)$ and $E(X^2)$. Indeed, we have

$$E(X) = \sum_{n=1}^{\infty} n(1-p)^{n-1}p = p\sum_{n=1}^{\infty} n(1-p)^{n-1} = p \cdot p^{-2} = p^{-1}$$

and

$$\begin{aligned}
E(X(X-1)) =& \sum_{n=1}^{\infty} n(n-1)(1-p)^{n-1}p \\
=& p(1-p)\sum_{n=1}^{\infty} n(n-1)(1-p)^{n-2} \\
=& p(1-p) \cdot (2p^{-3}) = 2p^{-2}(1-p).
\end{aligned}$$

Thus,

$$E(X^2) = E[X(X-1)] + E(X) = 2p^{-2}(1-p) + p^{-1} = (2-p)p^{-2}.$$

The variance is then given by

$$\text{Var}(X) = E(X^2) - (E(X))^2 = (2-p)p^{-2} - p^{-2} = \frac{1-p}{p^2}.$$

Next, observe that for $k = 1, 2, \cdots$ we have

$$\Pr(X \geq k) = \sum_{n=k}^{\infty} p(1-p)^{n-1} = p(1-p)^{k-1}\sum_{n=0}^{\infty}(1-p)^n = \frac{p(1-p)^{k-1}}{1-(1-p)} = (1-p)^{k-1}$$

and

$$\Pr(X \leq k) = 1 - \Pr(X \geq k+1) = 1 - (1-p)^k.$$

From this, one can find the cdf X given by

$$F(x) = \Pr(X \leq x) = \begin{cases} 0 & x < 1 \\ 1 - (1-p)^{\lfloor x \rfloor} & x \geq 1 \end{cases}$$

Example 22.2
Used watch batteries are tested one at a time until a good battery is found. Let X denote the number of batteries that need to be tested in order to find a good one. Find the expected value of X, given that $\Pr(X > 3) = 0.5$.

Solution.
X has geometric distribution, so $\Pr(X > 3) = \Pr(X \geq 4) = (1-p)^3$. Setting this equal to 1/2 and solving for p gives $p = 1 - 2^{-\frac{1}{3}}$. Therefore,

$$E(X) = \frac{1}{p} = \frac{1}{1 - 2^{-\frac{1}{3}}} = 4.847 \blacksquare$$

Example 22.3
From past experience it is noted that 3% of customers at an ATM machine make deposits on Sundays.
(a) What is the probability that the first deposit was made with the 5^{th} customer who used the ATM?
(b) What is the probability that the first deposit was made when 5 customers used the ATM?

Solution.
(a) Let X be the number of customers who used the ATM before the first deposit was made. Then X is a geometric random variable with $p = 0.03$. If 5 customers used the ATM before a deposit was made, then the first four customers did not make a deposit and the fifth customer made the deposit. Hence,

$$\Pr(X = 5) = 0.03(0.97)^4 = 0.027.$$

(b)

$$\Pr(X \leq 5) = 1 - 0.97^5 \approx 0.141 \blacksquare$$

Example 22.4
Assume that every time you attend your high school reunion there is a probability of 0.1 that your high school prom companion will not show up. Assume her arrival to any given reunion is independent of her arrival (or non-arrival) to any other reunion. What is the expected number of high school reunions you must attend until the time your prom miss the reunion?

Solution.
Let X be the number of reunions you must attend until you arrive to find your prom companion is absent, then X has a Geometric distribution with parameter $p = 0.1$. Thus

$$\Pr(X = n) = 0.1(1-p)^{n-1}, \quad n = 1, 2, \cdots$$

and

$$E(X) = \frac{1}{p} = 10 \blacksquare$$

Example 22.5
An archer shoots arrows at a circular target where the central portion of the target inside is called the bull. Suppose that an archer hits the target 70% of the time. Let X be the number of shoots until the first hit. Find the expected value and the standard deviation of X.

Solution.
X is a geometric random variable with $p = 0.7$. Thus, $E(X) = \frac{1}{p} = \frac{1}{0.7} = 0.43$ and $\sigma_X = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{0.3}{0.7^2}} = 0.78 \blacksquare$

Practice Problems

Problem 22.1
A box of candies contains 5 KitKat, 4 M&M, and 1 Crunch. Candies are drawn, with replacement, until a Crunch is found. If X is the random variable counting the number of trials until a Crunch appears, then
(a) What is the probability that the Crunch appears on the first trial?
(b) What is the probability that the Crunch appears on the second trial?
(c) What is is the probability that the Crunch appears on the n^{rmth} trial.

Problem 22.2
The probability that a computer chip is defective is 0.10. Each computer is checked for inspection as it is produced. Find the probability that at least 10 computer chips must be checked to find one that is defective.

Problem 22.3
Suppose a certain exam is classified as either difficult (with probability 90/92) or fair (with probability 2/92). Exams are taken one after the other. What is the probability that at least 4 difficult exams will occur before the first fair one?

Problem 22.4
Assume that every time you hot salsa, there is a 0.001 probability that you will get heartburn, independent of all other times you eaten hot salsa.
(a) What is the probability you will eat hot salsa two or less times until your first heartburn?
(b) What is the expected number of times you will eat hot salsa until you get your first heartburn?

Problem 22.5
Consider the experiment of flipping three coins simultaneously. Let a success be when the three outcomes are the same. What is the probability that
(a) exactly three rounds of flips are needed for the first success?
(b) more than four rounds are needed?

Problem 22.6
You roll a fair die repeatedly. Let a success be when the die shows either a 1 or a 3. Let X be the number of times you roll the die.
(a) What is $\Pr(X = 3)$? What is $\Pr(X = 50)$?
(b) Find the expected value of X.

Problem 22.7
Fifteen percent of the cars in a dealer's showroom have a sunroof. A salesperson starts showing you cars at random, one after the other. Let X be the number of cars with a sunroof that you see, before the first car that has no sunroof.
(a) What is the probability distribution function of X
(b) What is the probability distribution of $Y = X + 1$?

Problem 22.8
A study of car batteries shows that 3% of car batteries produced by a certain machine are defective. The batteries are put into packages of 20 batteries for distribution to retailers.
(a) What is the probability that a randomly selected package of batteries will contain at least 2 defective batteries?
(b) Suppose we continue to select packages of batteries randomly from the production site. What is the probability that it will take fewer than five packages to find a package with at least 2 defective batteries?

Problem 22.9
Show that the Geometric distribution with parameter p satisfies the equation

$$\Pr(X > i + j | X > i) = \Pr(X > j).$$

This says that the Geometric distribution satisfies the **memoryless property**

Problem 22.10 ‡
As part of the underwriting process for insurance, each prospective policyholder is tested for high blood pressure. Let X represent the number of tests completed when the first person with high blood pressure is found. The expected value of X is 12.5.
Calculate the probability that the sixth person tested is the first one with high blood pressure.

Problem 22.11
Suppose that the probability for an applicant to get a job offer after an interview is 0.1. An applicant plans to keep trying out for more interviews until she gets offered. Assume outcomes of interviews are independent.
(a) How many interviews does she expect to have to take in order to get a job offer?
(b) What is the probability she will need to attend more than 2 interviews?

Problem 22.12
In each of the following you are to determine whether the problem is a binomial type problem or a geometric type. In each case, find the probability mass function $p(x)$. Assume outcomes of individual trials are independent with constant probability of success.
(a) A arch shooter will aim at the target until one successfully hits it. The underlying probability of success is 0.40.
(b) A clinical trial enrolls 20 patients with a rare disease. Each patient is given an experimental therapy, and the number of patients showing marked improvement is observed. The true underlying probability of success is 0.60.

23 Negative Binomial Random Variable

The geometric distribution is the distribution of the number of bernoulli trials needed to get the first sucess. In this section we consider an extension of this distribution. We will study the distribution of the number of independent Bernoulli trials needed to get the r^{th} success.
Consider a Bernoulli experiment where a success occurs with probability p and a failure occurs with probability $q = 1 - p$. Assume that the experiment continues, that is the trials are performed, until the r^{th} success occurs. For example, in the rolling of a fair die, let a success be when the die shows a 5. We roll the die repeatedly until the fourth time the face 5 appears. In this case, $p = \frac{1}{6}$ and $r = 4$.
The random variable X, the number of trials needed to get the r^{th} success, has a **negative binomial** distribution with parameters r and p. It is worth mentioning the difference between the binomial distribution and the negative binomial distribution: In the binomial idistribution, X is the number of success in a fixed number of independent Bernoulli trials n. In the negative binomial distribution, X is the number of trials needed to get a fixed number of successes r.
For the r^{th} success to occur on the n^{th} trial, there must have been $r - 1$ successes among the first $n - 1$ trials. The number of ways of distributing $r - 1$ successes among $n - 1$ trials is ${}_{n-1}C_{r-1}$. But the probability of having $r - 1$ successes and $n - r$ failures is $p^{r-1}(1-p)^{n-r}$. The probability of the r^{th} success is p. Thus, the product of these three terms (using independence) is the probability that there are r successes and $n - r$ failures in the n trials, with the r^{th} success occurring on the n^{th} trial. Hence, the probability mass function of X is

$$p(n) = \Pr(X = n) = {}_{n-1}C_{r-1}p^r(1-p)^{n-r},$$

where $n = r, r + 1, \cdots$ (In order to have r successes there must be at least r trials.)
Note that if $r = 1$ then X is a geometric random variable with parameter p.
The negative binomial distribution is sometimes defined in terms of the random variable $Y =$ number of failures before the r^{th} success. This formulation is statistically equivalent to the one given above in terms of $X =$ number of trials at which the r^{th} success occurs, since $Y = X - r$. The alternative form of the negative binomial distribution is

$$\Pr(Y = y) = {}_{r+y-1}C_y p^r(1-p)^y = {}_{r+y-1}C_{r-1}p^r(1-p)^y, \quad y = 0, 1, 2, \cdots.$$

In this form, the negative binomial distribution is used when the number of successes is fixed and we are interested in the number of failures before reaching the fixed number of successes.
Note that the binomial coefficient

$$_{r+y-1}C_y = \frac{(r+y-1)!}{y!(r-1)!} = \frac{(y+r-1)(y+r-2)\cdots(r+1)r}{y!}$$

can be alternatively written in the following manner, expalining the name "negative binomial:

$$_{r+y-1}C_y = (-1)^y \frac{(-r)(-r-1)\cdots(-r-y+1)}{y!} = (-1)^y {}_yC_{-r}$$

which is the defining equation for binomial coefficient with negative integers.
Now, recalling the Taylor series expansion of the function $f(t) = (1-t)^{-r}$ at $t = 0$,

$$\begin{aligned}(1-t)^{-r} =& \sum_{k=0}^{\infty} (-1)^k {}_{-r}C_k t^k \\ =& \sum_{k=0}^{\infty} {}_{r+k-1}C_k t^k, \quad -1 < t < 1\end{aligned}$$

Thus,

$$\begin{aligned}\sum_{y=0}^{\infty} \Pr(Y=y) =& \sum_{y=0}^{\infty} {}_{r+y-1}C_y p^r (1-p)^y \\ =& p^r \sum_{y=0}^{\infty} {}_{r+y-1}C_y (1-p)^y \\ =& p^r \cdot p^{-r} = 1\end{aligned}$$

This shows that $p(y)$ is indeed a probability mass function.

Example 23.1
A research study is concerned with the side effects of a new drug. The drug is given to patients, one at a time, until two patients develop sided effects. If the probability of getting a side effect from the drug is $\frac{1}{6}$, what is the probability that eight patients are needed?

Solution.
Let Y be the number of patients who do not show side effects. Then Y follows a negative binomial distribution with $r = 2,\ y = 6$, and $p = \frac{1}{6}$. Thus,

$$\Pr(Y = 6) = {}_{2+6-1}C_6 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^6 \approx 0.0651 \blacksquare$$

Example 23.2
A person is conducting a phone survey. Define "success" as the event a person completes the survey and let Y be the number of failures before the third success. What is the probability that there are 10 failures before the third success? Assume that 1 out of 6 people contacted completed the survey.

Solution.
The probability that there are 10 failures before the third success is given by

$$\Pr(Y = 10) = {}_{3+10-1}C_{10} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^{10} \approx 0.0493 \blacksquare$$

Example 23.3
A four-sided die is rolled repeatedly. A success is when the die shows a 1. What is the probability that the tenth success occurs in the fortieth attempt?

Solution.
Let X number of attempts at which the tenth success occurs. Then X is a negative binomial random variable with parameters $r = 10$ and $p = 0.25$. Thus,

$$\Pr(X = 40) = {}_{40-1}C_{10-1}(0.25)^{10}(0.75)^{30} \approx 0.0360911 \blacksquare$$

Expected value and Variance

The expected value of Y is

$$\begin{aligned}
E(Y) =& \sum_{y=0}^{\infty} y_{r+y-1}C_y p^r(1-p)^y \\
=& \sum_{y=1}^{\infty} \frac{(r+y-1)!}{(y-1)!(r-1)!} p^r(1-p)^y \\
=& \sum_{y=1}^{\infty} \frac{r(1-p)}{p} {}_{r+y-1}C_{y-1} p^{r+1}(1-p)^{y-1} \\
=& \frac{r(1-p)}{p} \sum_{z=0}^{\infty} {}_{r+1+z-1}C_z p^{r+1}(1-p)^z \\
=& \frac{r(1-p)}{p}
\end{aligned}$$

It follows that

$$E(X) = E(Y+r) = E(Y) + r = \frac{r}{p}.$$

Similarly,

$$\begin{aligned}
E(Y^2) =& \sum_{y=0}^{\infty} y^2{}_{r+y-1}C_y p^r(1-p)^y \\
=& \frac{r(1-p)}{p} \sum_{y=1}^{\infty} y \frac{(r+y-1)!}{(y-1)!r!} p^{r+1}(1-p)^{y-1} \\
=& \frac{r(1-p)}{p} \sum_{z=0}^{\infty} (z+1)_{r+1+z-1}C_z p^{r+1}(1-p)^z \\
=& \frac{r(1-p)}{p}(E(Z)+1)
\end{aligned}$$

where Z is the negative binomial random variable with parameters $r+1$ and p. Using the formula for the expected value of a negative binomial random variable gives that

$$E(Z) = \frac{(r+1)(1-p)}{p}$$

Thus,

$$E(Y^2) = \frac{r^2(1-p)^2}{p^2} + \frac{r(1-p)}{p^2}.$$

The variance of Y is

$$\text{Var}(Y) = E(Y^2) - [E(Y)]^2 = \frac{r(1-p)}{p^2}.$$

Since $X = Y + r$,

$$\text{Var}(X) = \text{Var}(Y) = \frac{r(1-p)}{p^2}.$$

Example 23.4
A person is conducting a phone survey. Suppose that 1 of 6 people contacted will complete the survey. Define "success" as the event a person completes the survey and let Y be the number of failures before the third success. Find $E(Y)$ and $\text{Var}(Y)$.

Solution.
The expected value of Y is

$$E(Y) = \frac{r(1-p)}{p} = 15$$

and the variance is

$$\text{Var}(Y) = \frac{r(1-p)}{p^2} = 90 \blacksquare$$

Practice Problems

Problem 23.1
Consider a biased coin with the probability of getting heads is 0.1. Let X be the number of flips needed to get the 8^{th} heads.
(a) What is the probability of getting the 8^{th} heads on the 50^{th} toss?
(b) Find the expected value and the standard deviation of X.

Problem 23.2
Recently it is found that the bottom of the mediterranean sea near Cyprus has potential of oil discovery. Suppose that a well oil drilling has 20% chance of striking oil. Find the probability that the third oil strike comes on the 5^{th} well drilled.

Problem 23.3
Consider a 52-card deck. Repeatedly draw a card with replacement and record its face value. Let X be the number of trials needed to get three kings.
(a) What is the distribution of X? (b) What is the probability that $X = 39$?

Problem 23.4
Repeatdly roll a fair die until the outcome 3 has accurred for the 4^{th} time. Let X be the number of times needed in order to achieve this goal. Find $E(X)$ and $\text{Var}(X)$.

Problem 23.5
Find the probability of getting the fourth head on the ninth flip of a fair coin.

Problem 23.6
There is 75% chance to pass the written test for a driver's license. What is the probability that a person will pass the test on the second try?

Problem 23.7 ‡
A company takes out an insurance policy to cover accidents that occur at its manufacturing plant. The probability that one or more accidents will occur during any given month is $\frac{3}{5}$.
The number of accidents that occur in any given month is independent of the number of accidents that occur in all other months.
Calculate the probability that there will be at least four months in which no accidents occur before the fourth month in which at least one accident occurs.

Problem 23.8
Somehow waiters at a cafe are extremely distracted today and are mixing orders giving customers decaf coffee when they ordered regular coffee. Suppose that there is 60% chance of making such a mistake in the order. What is the probability of getting the second decaf on the seventh order of regular coffee?

Problem 23.9
A machine that produces computer chips produces 3 defective chips out of 100. Computer chips are delvered to retailers in packages of 20 chips each.
(a) A package is selected randomly. What is the probability that the package will contain at least 2 defective chips?
(b) What is the probability that the tenth package selected is the third to contain at least two defective chips?

Problem 23.10
Let X be a negative binomial distribution with $r = 2$ and $p = 0.1$. Find $E(X)$ and σ_X.

Problem 23.11
Suppose that the probability of a child exposed to the flu will catch the flu is 0.40. What is the probability that the tenth child exposed to the flu will be the third to catch it?

Problem 23.12
In rolling a fair die repeatedly (and independently on successive rolls), find the probability of getting the third "3" on the n^{th} roll.

Problem 23.13 ‡
Each time a hurricane arrives, a new home has a 0.4 probability of experiencing damage. The occurrences of damage in different hurricanes are independent. Calculate the mode of the number of hurricanes it takes for the home to experience damage from two hurricanes. Hint: The mode of X is the number that maximizes the probability mass function of X.

24 Hypergeometric Random Variable

Suppose we have a population of N objects which are divided into two types: Type A and Type B. There are n objects of Type A and $N-n$ objects of Type B. For example, a standard deck of 52 playing cards can be divided in many ways. Type A could be "Hearts" and Type B could be "All Others." Then there are 13 Hearts and 39 others in this population of 52 cards.
Suppose a random sample of size r is taken (without replacement) from the entire population of N objects. The **Hypergeometric random variable** X counts the total number of objects of Type A in the sample.
If $r \leq n$ then there could be at most r objects of Type A in the sample. If $r > n$, then there can be at most n objects of Type A in the sample. Thus, the value $\min\{r, n\}$ is the maximum possible number of objects of Type A in the sample.
On the other hand, if $r \leq N-n$, then all objects chosen may be of Type B. But if $r > N-n$, then there must be at least $r-(N-n)$ objects of Type A chosen. Thus, the value $\max\{0, r-(N-n)\}$ is the least possible number of objects of Type A in the sample.
What is the probability of having exactly k objects of Type A in the sample, where $\max\{0, r-(N-n)\} \leq k \leq \min\{r, n\}$? This is a type of problem that we have done before: In a group of N people there are n men (and the rest women). If we appoint a committee of r persons from this group at random, what is the probability there are exactly k men on it? The number of susbets of the group with cardinality r is ${}_NC_r$. The number of subsets of the men with cardinality k is ${}_nC_k$ and the number of subsets of the women with cardinality $r-k$ is ${}_{N-n}C_{r-k}$. Thus, the probability of getting exactly k men on the committee is

$$p(k) = \Pr(X = k) = \frac{{}_nC_k \cdot {}_{N-n}C_{r-k}}{{}_NC_r}, \ k = 0, 1, \cdots, r.$$

This is the probability mass function of X. Note that $p(k) \geq 0$ and

$$\sum_{k=0}^{r} \frac{{}_nC_k \cdot {}_{N-n}C_{r-k}}{{}_NC_r} = 1$$

The proof of this result follows from

Theorem 24.1 *(***Vendermonde's identity***)*

$$_{n+m}C_r = \sum_{k=0}^{r} {}_nC_k \cdot {}_mC_{r-k}.$$

Proof.
Suppose a committee consists of n men and m women. In how many ways can a subcommittee of r members be formed? The answer is $_{n+m}C_r$. But on the other hand, the answer is the sum over all possible values of k, of the number of subcommittees consisting of k men and $r-k$ women ■

Example 24.1
An urn contains 70 red marbles and 30 green marbles. If we draw out 20 without replacement, what is the probability of getting exactly 14 red marbles?

Solution.
If X is the number of red marbles, then X is a hypergeometric random variable with parameters $N = 100, r = 20, n = 70$. Thus,

$$\Pr(X = 14) = \frac{_{70}C_{14} \cdot {}_{30}C_6}{_{100}C_{20}} \approx 0.21 ■$$

Example 24.2
A barn consists of 13 cows, 12 pigs and 8 horses. A group of 8 will be selected to participate in the city fair. What is the probability that exactly 5 of the group will be cows?

Solution.
Let X be the number of cows in the group. Then X is hypergeometric random variable with parameters $N = 33, r = 8, n = 13$. Thus,

$$\Pr(X = 5) = \frac{_{13}C_5 \cdot {}_{20}C_3}{_{33}C_8} \approx 0.10567 ■$$

Next, we find the expected value of a hypergeometric random variable with parameters N, n, r. Let k be a positive integer. Then

$$\begin{aligned} E(X^k) &= \sum_{i=0}^{r} i^k \Pr(X = i) \\ &= \sum_{i=0}^{r} i^k \frac{_nC_i \cdot {}_{N-n}C_{r-i}}{_NC_r} \end{aligned}$$

Using the identities

$$i_nC_i = n_{n-1}C_{i-1} \quad \text{and} \quad r_NC_r = N_{N-1}C_{r-1}$$

we obtain that

$$\begin{aligned}
E(X^k) =& \frac{nr}{N}\sum_{i=1}^{r} i^{k-1}\frac{{}_{n-1}C_{i-1}\cdot {}_{N-n}C_{r-i}}{{}_{N-1}C_{r-1}}\\
=& \frac{nr}{N}\sum_{j=0}^{r-1}(j+1)^{k-1}\frac{{}_{n-1}C_j\cdot {}_{(N-1)-(n-1)}C_{(r-1)-j}}{{}_{N-1}C_{r-1}}\\
=& \frac{nr}{N}E[(Y+1)^{k-1}]
\end{aligned}$$

where Y is a hypergeometric random variable with paramters $N-1, n-1,$ and $r-1$. By taking $k=1$ we find

$$E(X) = \frac{nr}{N}.$$

Now, by setting $k=2$ we find

$$E(X^2) = \frac{nr}{N}E(Y+1) = \frac{nr}{N}\left(\frac{(n-1)(r-1)}{N-1}+1\right).$$

Hence,

$$\begin{aligned}
Var(X) =& E(X^2) - [E(X)]^2 = \frac{nr}{N}\left[\frac{(n-1)(r-1)}{N-1}+1-\frac{nr}{N}\right]\\
=& \frac{nr}{N}\cdot\frac{N-r}{N}\cdot\frac{N-n}{N-1}.
\end{aligned}$$

Example 24.3
The faculty senate of a certain college has 20 members. Suppose there are 12 men and 8 women. A committee of 10 senators is selected at random.
(a) What is the probability that there will be 6 men and 4 women on the committee?
(b) What is the expected number of men on this committee?
(c) What is the variance of the number of men on this committee?

Solution.
Let X be the number of men of the committee of 10 selected at random. Then X is a hypergeometric random variable with $N = 20, r = 10,$ and

$n = 12.$.
(a) $\Pr(X = 6) = \frac{{}_{12}C_6 \cdot {}_8C_4}{{}_{20}C_{10}}$.
(b) $E(X) = \frac{nr}{N} = \frac{12(10)}{20} = 6$
(c) $\text{Var}(X) = n \cdot \frac{r}{N} \cdot \frac{N-r}{N} \cdot \frac{N-n}{N-1} = 1.263$ ■

Example 24.4
A package of 15 computer chips contains 6 defective chips and 9 nondefective. Five chips are randomly selected without replacement.
(a) What is the probability that there are 2 defective and 3 nondefective chips in the sample?
(b) What is the probability that there are at least 3 nondefective chips in the sample?
(c) What is the expected number of defective chips in the sample?

Solution.
(a) Let X be the number of defective chips in the sample. Then, X has a hypergeometric distribution with $n = 6, N = 15, r = 5$. The desired probability is

$$\Pr(X = 2) = \frac{{}_6C_2 \cdot {}_9C_3}{{}_{15}C_5} = \frac{420}{1001}$$

(b) Note that the event that there are at least 3 nondefective chips in the sample is equivalent to the event that there are at most 2 defective chips in the sample, i.e. $\{X \leq 2\}$. So, we have

$$\begin{aligned}
\Pr(X \leq 2) =&\Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2)\\
=&\frac{{}_6C_0 \cdot {}_9C_5}{{}_{15}C_5} + \frac{{}_6C_1 \cdot {}_9C_4}{{}_{15}C_5} + \frac{{}_6C_2 \cdot {}_9C_3}{{}_{15}C_5}\\
=&\frac{714}{1001}
\end{aligned}$$

(c) $E(X) = \frac{rn}{N} = 5 \cdot \frac{6}{15} = 2$ ■

Practice Problems

Problem 24.1
Five cards are drawn randomly and without replacement from a deck of 52 playing cards. Find the probability of getting exactly two black cards.

Problem 24.2
An urn contains 15 red marbles and 10 blue ones. Seven marbles were randomly drawn without replacement. Find the probability of picking exactly 3 red marbles.

Problem 24.3
A lottery game consists of matching 6 numbers from the official six drawn numbers out of 53 numbers. Let X equal the number of matches. Find the probability distribution function.

Problem 24.4
A package of 20 computer chips contains 4 defective chips. Randomly select 10 chips without replacement. Compute the probability of obtaining exactly 3 defective chips.

Problem 24.5
A wallet contains 10 \$50 bills and 190 \$1 bills. You randomly choose 10 bills without replacement. What is the probability that you will choose exactly 2 \$50 bills?

Problem 24.6
A batch of 8 components contains 2 defective components and 6 good ones. Randomly select four components without replacement.
(a) What is the probability that all four components are good?
(b) What are the mean and variance for the number of good components?

Problem 24.7
In Texas all vehicles are subject to annual inspection. A transportation company has a fleet of 20 trucks in which 7 do not meet the standards for passing inspection. Five trucks are randomly selected for inspection. What is the probability of no more than 2 trucks that fail to have the standards for passing inspection being selected?

Problem 24.8
A recent study shows that in a certain city 2,477 cars out of 123,850 are stolen. The city police are trying to find the stolen cars. Suppose that 100 randomly chosen cars are checked by the police. Find the expression that gives the probability that exactly 3 of the chosen cars are stolen. You do not need to give the numerical value of this expression.

Problem 24.9
Consider a suitcase with 7 shirts and 3 pants. Suppose we draw 4 items without replacement from the suitcase. Let X be the total number of shirts we get. Compute $\Pr(X \leq 1)$.

Problem 24.10
A group consists of 4 women and 20 men. A committe of six is to be formed. Using the appropriate hypergeometric distribution, what is the probability that none of the women are on the committee?

Problem 24.11
A jar contains 10 white balls and 15 black balls. Let X denote the number of white balls in a sample of 10 balls selected at random and without replacement. Find $\frac{\text{Var}(X)}{E(X)}$.

Problem 24.12
Among the 48 applicants for an actuarial position, 30 have a college degree in actuarial science. Ten of the applicants are randomly chosen for interviews. Let X be the number of applicants among these ten who have a college degree in actuarial science. Find $\Pr(X \leq 8)$.

Cumulative and Survival Distribution Functions

In this chapter, we study the properties of two important functions in probability theory related to random variables: the cumulative distribution function and the survival distribution function.

25 The Cumulative Distribution Function

In this section, we will discuss properties of the cumulative distribution function that are valid to a random variable of type discrete, continuous or mixed. Recall from Section 14 that if X is a random variable then the **cumulative distribution function** (abbreviated c.d.f) is the function

$$F(t) = \Pr(X \leq t).$$

First, we prove that probability is a continuous set function. In order to do that, we need the following definitions.
A sequence of sets $\{E_n\}_{n=1}^{\infty}$ is said to be **increasing** if

$$E_1 \subset E_2 \subset \cdots \subset E_n \subset E_{n+1} \subset \cdots$$

whereas it is said to be a **decreasing** sequence if

$$E_1 \supset E_2 \supset \cdots \supset E_n \supset E_{n+1} \supset \cdots$$

If $\{E_n\}_{n=1}^{\infty}$ is an increasing sequence of events we define a new event

$$\lim_{n\to\infty} E_n = \bigcup_{n=1}^{\infty} E_n.$$

For a decreasing sequence we define

$$\lim_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} E_n.$$

Proposition 25.1
If $\{E_n\}_{n\geq 1}$ is either an increasing or decreasing sequence of events then
(a)

$$\lim_{n\to\infty} \Pr(E_n) = \Pr(\lim_{n\to\infty} E_n)$$

that is

$$\Pr\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \Pr(E_n)$$

for an increasing sequence and
(b)

$$\Pr\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \Pr(E_n)$$

for a decreasing sequence.

Proof.
(a) Suppose first that $E_n \subset E_{n+1}$ for all $n \geq 1$. Define the events

$$\begin{aligned} F_1 =& E_1 \\ F_n =& E_n \cap E_{n-1}^c, \quad n > 1 \end{aligned}$$

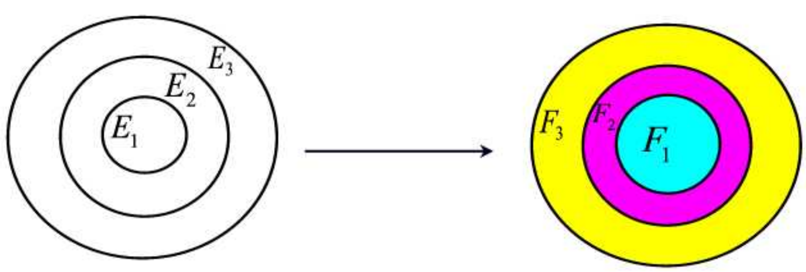


Figure 25.1

These events are shown in the Venn diagram of Figure 25.1. Note that for $n > 1$, F_n consists of those outcomes in E_n that are not in any of the earlier

$E_i,\ i < n$. Clearly, for $i \neq j$ we have $F_i \cap F_j = \emptyset$. Also, $\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} E_n$ and for $n \geq 1$ we have $\bigcup_{i=1}^{n} F_i = \bigcup_{i=1}^{n} E_i$. From these properties we have

$$\begin{aligned}
\Pr(\lim_{n\to\infty} E_n) =& \Pr(\bigcup_{n=1}^{\infty} E_n)\\
=& \Pr(\bigcup_{n=1}^{\infty} F_n)\\
=& \sum_{n=1}^{\infty} \Pr(F_n)\\
=& \lim_{n\to\infty} \sum_{i=1}^{n} \Pr(F_i)\\
=& \lim_{n\to\infty} \Pr(\bigcup_{i=1}^{n} F_i)\\
=& \lim_{n\to\infty} \Pr(\bigcup_{i=1}^{n} E_i)\\
=& \lim_{n\to\infty} \Pr(E_n)
\end{aligned}$$

(b) Now suppose that $\{E_n\}_{n\geq 1}$ is a decreasing sequence of events. Then $\{E_n^c\}_{n\geq 1}$ is an increasing sequence of events. Hence, from part (a) we have

$$\Pr(\bigcup_{n=1}^{\infty} E_n^c) = \lim_{n\to\infty} \Pr(E_n^c)$$

By De Morgan's Law we have $\bigcup_{n=1}^{\infty} E_n^c = \left(\bigcap_{n=1}^{\infty} E_n\right)^c$. Thus,

$$\Pr\left(\left(\bigcap_{n=1}^{\infty} E_n\right)^c\right) = \lim_{n\to\infty} \Pr(E_n^c).$$

Equivalently,

$$1 - \Pr\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} [1 - \Pr(E_n)] = 1 - \lim_{n\to\infty} \Pr(E_n)$$

or

$$\Pr\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \Pr(E_n) \ \blacksquare$$

Proposition 25.2
F is a nondecreasing function; that is, if $a < b$ then $F(a) \leq F(b)$.

Proof.
Suppose that $a < b$. Then $\{s : X(s) \leq a\} \subseteq \{s : X(s) \leq b\}$. This implies that $\Pr(X \leq a) \leq \Pr(X \leq b)$. Hence, $F(a) \leq F(b)$ ■

Example 25.1
Determine whether the given values can serve as the values of a distribution function of a random variable with the range $x = 1, 2, 3, 4$.

$$F(1) = 0.5, \;\; F(2) = 0.4, \;\; F(3) = 0.7, \text{ and } F(4) = 1.0$$

Solution.
Since $F(2) < F(1)$, F violates is not increasing and therefore F can not be a cdf ■

Proposition 25.3
F is continuous from the right. That is, $\lim_{t\to b^+} F(t) = F(b)$.

Proof.
Let $\{b_n\}$ be a decreasing sequence that converges to b with $b_n \geq b$ for all n. Define $E_n = \{s : X(s) \leq b_n\}$. Then $\{E_n\}_{n\geq 1}$ is a decreasing sequence of events such that $\bigcap_{n=1}^{\infty} E_n = \{s : X(s) \leq b\}$. By Proposition 25.1 we have

$$\lim_{n\to\infty} F(b_n) = \lim_{n\to\infty} \Pr(E_n) = \Pr\left(\bigcap_{n=1}^{\infty} E_n\right) = \Pr(X \leq b) = F(b) \; ■$$

Proposition 25.4
(a) $\lim_{b\to-\infty} F(b) = 0$
(b) $\lim_{b\to\infty} F(b) = 1$

Proof.
(a) Note that $\lim\limits_{x\to-\infty} F(x) = \lim\limits_{n\to\infty} F(x_n)$ where (x_n) is a decreasing sequence such that $x_n \to -\infty$. Define $E_n = \{s \in S : X(s) \leq x_n\}$. Then we have the nested chain $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$. Moreover,

$$\emptyset = \bigcap_{n=1}^{\infty} E_n.$$

By Proposition 25.1, we find

$$\lim_{x\to\infty} F(x) = \lim_{n\to-\infty} F(x_n) = \lim_{n\to\infty} \Pr(E_n) = \Pr\left(\bigcap_{n=1}^{\infty} E_n\right) = \Pr(\emptyset) = 0.$$

(b) Note that $\lim_{x\to\infty} F(x) = \lim_{n\to\infty} F(x_n)$ where (x_n) is an increasing sequence such that $x_n \to \infty$. Define $E_n = \{s \in S : X(s) \leq x_n\}$. Then we have the nested chain $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$. Moreover,

$$S = \bigcup_{n=1}^{\infty} E_n.$$

By Proposition 25.1, we find

$$\lim_{x\to\infty} F(x) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \Pr(E_n) = \Pr\left(\bigcup_{n=1}^{\infty} E_n\right) = \Pr(S) = 1 \blacksquare$$

Example 25.2
Determine whether the given values can serve as the values of a distribution function of a random variable with the range $x = 1, 2, 3, 4$.

$$F(1) = 0.3,\ F(2) = 0.5,\ F(3) = 0.8, \text{ and } F(4) = 1.2$$

Solution.
No because $F(4)$ exceeds 1 $\blacksquare$

All probability questions can be answered in terms of the c.d.f.

Proposition 25.5
For any random variable X and any real number a, we have

$$\Pr(X > a) = 1 - F(a).$$

Proof.
Let $A = \{x \in S : X(s) \leq a\}$. Then $A^c = \{s \in S : X(s) > a\}$. We have $\Pr(X > a) = \Pr(A^c) = 1 - \Pr(A) = 1 - \Pr(X \leq a) = 1 - F(a)$ $\blacksquare$

Example 25.3
Let X have probability mass function (pmf) $\Pr(x) = \frac{1}{8}$ for $x = 1, 2, \cdots, 8$. Find
(a) the cumulative distribution function (cdf) of X;
(b) $\Pr(X > 5)$.

Solution.
(a) The cdf is given by

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{\lfloor x \rfloor}{8} & 1 \le x \le 8 \\ 1 & x > 8 \end{cases}$$

where $[x]$ is the floor function of x.
(b) We have $\Pr(X > 5) = 1 - F(5) = 1 - \frac{5}{8} = \frac{3}{8}$ ■

Proposition 25.6
For any random variable X and any real number a, we have

$$\Pr(X < a) = \lim_{n \to \infty} F\left(a - \frac{1}{n}\right) = F(a^-).$$

Proof.
For each positive integer n, define $E_n = \{s \in S : X(s) \le a - \frac{1}{n}\}$. Then $\{E_n\}$ is an increasing sequence of sets such that

$$\bigcup_{n=1}^{\infty} E_n = \{s \in S : X(s) < a\}.$$

We have

$$\begin{aligned} \Pr(X < a) =& \Pr\left(\lim_{n \to \infty}\left\{X \le a - \frac{1}{n}\right\}\right) \\ =& \lim_{n \to \infty} \Pr\left(X \le a - \frac{1}{n}\right) \\ =& \lim_{n \to \infty} F\left(a - \frac{1}{n}\right) = F(a^-) \ ■ \end{aligned}$$

Note that $\Pr(X < a)$ does not necessarily equal $F(a)$, since $F(a)$ also includes the probability that X equals a.

Corollary 25.1

$$\Pr(X \ge a) = 1 - \lim_{n \to \infty} F\left(a - \frac{1}{n}\right) = 1 - F(a^-).$$

Proposition 25.7
If $a < b$ then $\Pr(a < X \leq b) = F(b) - F(a)$.

Proof.
Let $A = \{s : X(s) > a\}$ and $B = \{s : X(s) \leq b\}$. Note that $\Pr(A \cup B) = 1$. Then

$$\begin{aligned}\Pr(a < X \leq b) =&\Pr(A \cap B)\\ =&\Pr(A) + \Pr(B) - \Pr(A \cup B)\\ =&(1 - F(a)) + F(b) - 1 = F(b) - F(a) \blacksquare\end{aligned}$$

Proposition 25.8
If $a < b$ then $\Pr(a \leq X < b) = F(b^-) - F(a^-)$.

Proof.
Let $A = \{s : X(s) \geq a\}$ and $B = \{s : X(s) < b\}$. Note that $\Pr(A \cup B) = 1$. We have,

$$\begin{aligned}\Pr(a \leq X < b) =&\Pr(A \cap B)\\ =&\Pr(A) + \Pr(B) - \Pr(A \cup B)\\ =&\left(1 - \lim_{n\to\infty} F\left(a - \frac{1}{n}\right)\right) + \lim_{n\to\infty} F\left(b - \frac{1}{n}\right) - 1\\ =&\lim_{n\to\infty} F\left(b - \frac{1}{n}\right) - \lim_{n\to\infty} F\left(a - \frac{1}{n}\right)\\ =&F(b^-) - F(a^-) \blacksquare\end{aligned}$$

Proposition 25.9
If $a < b$ then $\Pr(a \leq X \leq b) = F(b) - \lim_{n\to\infty} F\left(a - \frac{1}{n}\right) = F(b) - F(a^-)$.

Proof.
Let $A = \{s : X(s) \geq a\}$ and $B = \{s : X(s) \leq b\}$. Note that $\Pr(A \cup B) = 1$. Then

$$\begin{aligned}\Pr(a \leq X \leq b) =&\Pr(A \cap B)\\ =&\Pr(A) + \Pr(B) - \Pr(A \cup B)\\ =&\left(1 - \lim_{n\to\infty} F\left(a - \frac{1}{n}\right)\right) + F(b) - 1\\ =&F(b) - F(a^-) \blacksquare\end{aligned}$$

Example 25.4
Show that $\Pr(X = a) = F(a) - F(a^-)$.

Solution.
Applying the previous result we can write

$$\Pr(X = a) = \Pr(a \le x \le a) = F(a) - F(a^-) \blacksquare$$

Proposition 25.10
If $a < b$ then $\Pr(a < X < b) = F(b^-) - F(a)$.

Proof.
Let $A = \{s : X(s) > a\}$ and $B = \{s : X(s) < b\}$. Note that $\Pr(A \cup B) = 1$. Then

$$\begin{aligned}
\Pr(a < X < b) =&\Pr(A \cap B)\\
=&\Pr(A) + \Pr(B) - \Pr(A \cup B)\\
=&(1 - F(a)) + \lim_{n\to\infty} F\left(b - \frac{1}{n}\right) - 1\\
=&\lim_{n\to\infty} F\left(b - \frac{1}{n}\right) - F(a)\\
=&F(b^-) - F(a) \blacksquare
\end{aligned}$$

Figure 25.2 illustrates a typical F for a discrete random variable X. Note that for a discrete random variable the cumulative distribution function will always be a step function with jumps at each value of x that has probability greater than 0 and the size of the step at any of the values $x_1, x_2, x_3, \cdots$ is equal to the probability that X assumes that particular value.

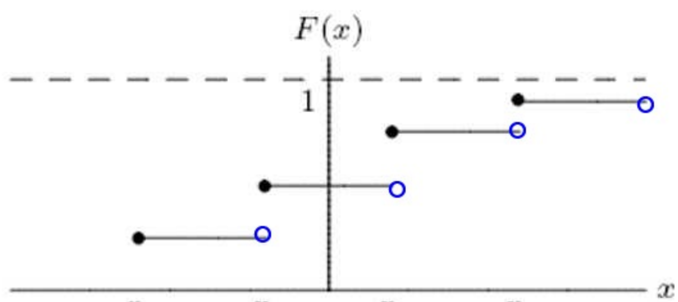


Figure 25.2

Example 25.5 *(*Mixed RV*)*
The distribution function of a random variable X, is given by

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{2}, & 0 \le x < 1 \\ \frac{2}{3}, & 1 \le x < 2 \\ \frac{11}{12}, & 2 \le x < 3 \\ 1, & 3 \le x \end{cases}$$

(a) Graph $F(x)$.
(b) Compute $\Pr(X < 3)$.
(c) Compute $\Pr(X = 1)$.
(d) Compute $\Pr(X > \frac{1}{2})$
(e) Compute $\Pr(2 < X \le 4)$.

Solution.
(a) The graph is given in Figure 25.3.
(b) $\Pr(X < 3) = \lim_{n\to\infty} \Pr\left(\left\{X \le 3 - \frac{1}{n}\right\}\right) = \lim_{n\to\infty} F\left(3 - \frac{1}{n}\right) = \frac{11}{12}$.
(c) $\Pr(X = 1) = \Pr(X \le 1) - \Pr(X < 1) = F(1) - \lim_{n\to\infty} F\left(1 - \frac{1}{n}\right) = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$.
(d) $\Pr(X > \frac{1}{2}) = 1 - \Pr(X \le \frac{1}{2}) = 1 - F(\frac{1}{2}) = \frac{3}{4}$.
(e) $\Pr(2 < X \le 4) = F(4) - F(2) = \frac{1}{12}$ ■

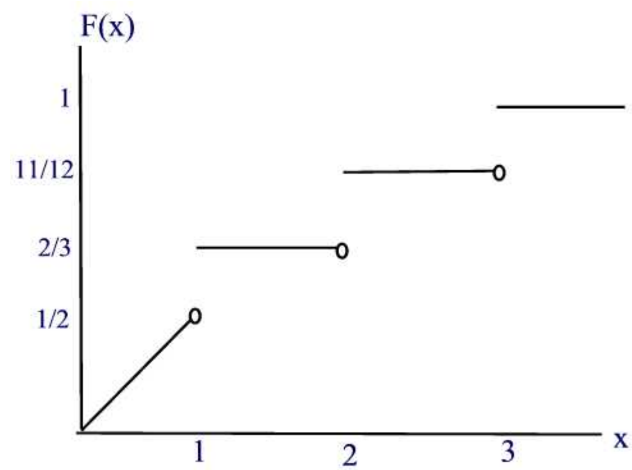


Figure 25.3

Example 25.6
If X has the cdf

$$F(x) = \begin{cases} 0, & x < -1 \\ \frac{1}{4}, & -1 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x < 3 \\ \frac{3}{4}, & 3 \leq x < 5 \\ 1, & x \geq 5. \end{cases}$$

find
(a) $\Pr(X \leq 3)$
(b) $\Pr(X = 3)$
(c) $\Pr(X < 3)$
(d) $\Pr(X \geq 1)$
(e) $\Pr(-0.4 < X < 4)$
(f) $\Pr(-0.4 \leq X < 4)$
(g) $\Pr(-0.4 < X \leq 4)$
(h) $\Pr(-0.4 \leq X \leq 4)$
(i) $\Pr(X = 5)$.

Solution.
(a) $\Pr(X \leq 3) = F(3) = \frac{3}{4}$.
(b) $\Pr(X = 3) = F(3) - F(3^-) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$
(c) $\Pr(X < 3) = F(3^-) = \frac{1}{2}$
(d) $\Pr(X \geq 1) = 1 - F(1^-) = 1 - \frac{1}{4} = \frac{3}{4}$
(e) $\Pr(-0.4 < X < 4) = F(4^-) - F(-0.4) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$
(f) $\Pr(-0.4 \leq X < 4) = F(4^-) - F(-0.4^-) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$
(g) $\Pr(-0.4 < X \leq 4) = F(4) - F(-0.4) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$
(h) $\Pr(-0.4 \leq X \leq 4) = F(4) - F(-0.4^-) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$
(i) $\Pr(X = 5) = F(5) - F(5^-) = 1 - \frac{3}{4} = \frac{1}{4}$ ■

Practice Problems

Problem 25.1
In your pocket, you have 1 dime, 2 nickels, and 2 pennies. You select 2 coins at random (without replacement). Let X represent the amount (in cents) that you select from your pocket.
(a) Give (explicitly) the probability mass function for X.
(b) Give (explicitly) the cdf, $F(x)$, for X.
(c) How much money do you expect to draw from your pocket?

Problem 25.2
We are inspecting a lot of 25 batteries which contains 5 defective batteries. We randomly choose 3 batteries. Let X = the number of defective batteries found in a sample of 3. Give the cumulative distribution function as a table.

Problem 25.3
Suppose that the cumulative distribution function is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{4} & 0 \le x < 1 \\ \frac{1}{2} + \frac{x-1}{4} & 1 \le x < 2 \\ \frac{11}{12} & 2 \le x < 3 \\ 1 & 3 \le x \end{cases}$$

(a) Find $\Pr(X = i), i = 1, 2, 3$.
(b) Find $\Pr(\frac{1}{2} < X < \frac{3}{2})$.

Problem 25.4
If the cumulative distribution function is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2} & 0 \le x < 1 \\ \frac{3}{5} & 1 \le x < 2 \\ \frac{4}{5} & 2 \le x < 3 \\ \frac{9}{10} & 3 \le x < 3.5 \\ 1 & 3.5 \le x \end{cases}$$

Calculate the probability mass function.

Problem 25.5
Consider a random variable X whose distribution function (cdf) is given by

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.1 & -2 \leq x < 1.1 \\ 0.3 & 1.1 \leq x < 2 \\ 0.6 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

(a) Give the probability mass function, $\Pr(x)$, of X, explicitly.
(b) Compute $\Pr(2 < X < 3)$.
(c) Compute $\Pr(X \geq 3)$.
(d) Compute $\Pr(X \geq 3 | X \geq 0)$.

Problem 25.6
Consider a random variable X whose probability mass function is given by

$$\Pr(x) = \begin{cases} p & x = -1.9 \\ 0.1 & x = -0.1 \\ 0.3 & x = 20p \\ p & x = 3 \\ 4p & x = 4 \\ 0 & \text{otherwise} \end{cases}$$

(a) What is p?
(b) Find $F(x)$ and sketch its graph.
(c) What is $F(0)$? What is $F(2)$? What is $F(F(3.1))$?
(d) What is $\Pr(2X - 3 \leq 4 | X \geq 2.0)$?
(e) Compute $E(F(X))$.

Problem 25.7
The cdf of X is given by

$$F(x) = \begin{cases} 0 & x < -4 \\ 0.3 & -4 \leq x < 1 \\ 0.7 & 1 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

(a) Find the probability mass function.
(b) Find the variance and the standard deviation of X.

Problem 25.8
In the game of "dice-flip", each player flips a coin and rolls one die. If the coin comes up tails, his score is the number of dots showing on the die. If the coin comes up heads, his score is twice the number of dots on the die. (i.e., (tails,4) is worth 4 points, while (heads,3) is worth 6 points.) Let X be the first player's score.
(a) Find the probability mass function $\Pr(x)$.
(b) Compute the cdf $F(x)$ for all numbers x.
(c) Find the probability that $X < 4$. Is this the same as $F(4)$?

Problem 25.9
A random variable X has cumulative distribution function

$$F(x) = \begin{cases} 0 & x \le 0 \\ \frac{x^2}{4} & 0 \le x < 1 \\ \frac{1+x}{4} & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

(a) What is the probability that $X = 0$? What is the probability that $X = 1$? What is the probability that $X = 2$?
(b) What is the probability that $\frac{1}{2} < X \le 1$?
(c) What is the probability that $\frac{1}{2} \le X < 1$?
(d) What is the probability that $X > 1.5$?
Hint: You may find it helpful to sketch this function.

Problem 25.10
Let X be a random variable with the following cumulative distribution function:

$$F(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \le x < \frac{1}{2} \\ \alpha & x = \frac{1}{2} \\ 1 - 2^{-2x} & x > \frac{1}{2} \end{cases}$$

(a) Find $\Pr(X > \frac{3}{2})$.
(b) Find $\Pr(\frac{1}{4} < X \le \frac{3}{4})$.
(c) Find α.
(d) Find $\Pr(X = \frac{1}{2})$.
(e) Sketch the graph of $F(x)$.

26 The Survival Distribution Function

Another key function describing a random variable is the survival distribution function.
The **survival function** (abbreviated SDF), also known as a **reliability function** is a property of any random variable that maps a set of events, usually associated with mortality or failure of some system, onto time. It captures the probability that the system will survive beyond a specified time. Thus, we define the survival distribution function by

$$S(x) = \Pr(X > x) = 1 - F(x).$$

It follows from the properties of the cumulative distribution function $F(x)$, that any random variable satisfies the properties: $S(-\infty) = 1$, $S(\infty) = 0$, $S(x)$ is right-continuous, and that $S(x)$ is nonincreasing. These four conditions are necessary and sufficient so that any nonnegative function $S(x)$ that satisfies these conditions serves as a survival function.

Remark 26.1
If X is the random variable representing the age of an individual when death occurs, i.e., refer to X as the **age-at-death** then $S(x)$ is the probability that an indiviual survived past the age of x. Thus, the term "survival".

Remark 26.2
For a discrete random variable, the survival function need not be left-continuous, that is, it is possible for its graph to jump down. When it jumps, the value is assigned to the bottom of the jump.

Example 26.1
Let X be a continuous random variable with survival distribution defined by $S(x) = e^{-0.34x}$ for $x \geq 0$ and 1 otherwise. Compute $\Pr(5 < X \leq 10)$.

Solution.
We have

$$\Pr(5 < X \leq 10) = F(10) - F(5) = S(5) - S(10) = e^{-0.34\times 5} - e^{-0.34\times 10} = 0.149 \blacksquare$$

Example 26.2
Let X be the random variable representing the age of death of an individual.

The survival distribution function for an individual is determined to be

$$S(x) = \begin{cases} 1, & x < 0 \\ \frac{75-x}{75}, & 0 \le x \le 75 \\ 0, & x > 75. \end{cases}$$

(a) Find the probability that the person dies before reaching the age of 18.
(b) Find the probability that the person lives more than 55 years.
(c) Find the probability that the person dies between the ages of 25 and 70.

Solution.
(a) We have

$$\Pr(X < 18) = \Pr(X \le 18) = F(18) = 1 - S(18) = 0.24.$$

(b) We have

$$\Pr(X > 55) = S(55) = 0.267.$$

(c) We have

$$\Pr(25 < X < 70) = F(70) - F(25) = S(25) - S(70) = 0.60 \blacksquare$$

Practice Problems

Problem 26.1
Consider the random variable X with survival distribution defined by

$$S(x) = \begin{cases} 1, & x < 0 \\ \frac{1}{10}(100-x)^{\frac{1}{2}}, & 0 \le x \le 100 \\ 0, & x > 100. \end{cases}$$

(a) Find the corresponding expression for the cumulative probability function.
(b) Compute $\Pr(65 < X \le 75)$.

Problem 26.2
Let X denote the age at death of an individual. The survival distribution is given by

$$S(x) = \begin{cases} 1, & x < 0 \\ 1 - \frac{x}{100}, & 0 \le x \le 100 \\ 0, & x > 100. \end{cases}$$

(a) Find the probability that a person dies before reaching the age of 30.
(b) Find the probability that a person lives more than 70 years.

Problem 26.3
If X is a continuous random variable then the survival distribution function is defined by

$$S(x) = \int_x^\infty f(t)dt$$

where $f(t)$ is called the **probability density function** of X. Show that $F'(x) = f(x)$.

Problem 26.4
Let X be a continuous random variable with cumulative distribution function

$$F(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-\lambda x}, & x \ge 0 \end{cases}$$

where $\lambda > 0$. Find the probability density function $f(x)$.

Problem 26.5
Given the cumulative distribution function

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

Find $S(x)$.

Calculus Prerequisite

In this chapter, we collect concepts from calculus deemed necessary in the understanding of the topics that follow later in this book.

27 Graphing Systems of Linear Inequalities in Two Variables

When evaluating double integrals over a certain region, the region under consideration is the solution to a system of linear inequalities in two variables. The purpose of this section is to represent the solution graphically.
By a **linear inequality** in the variables x and y we mean anyone of the following

$$ax + by \leq c, \; ax + by \geq c, \; ax + by < c, \; ax + by > c.$$

A pair of numbers (x, y) that satisfies a linear inequality is called a **solution.** A solution set of a linear inquality is a half-plane in the Cartesian coordinates system. The **boundary** of the region is graph of the line $ax + by = c$. The boundary is represented by a dashed line in the case of inequalities involving either $<$ or $>$. Otherwise, the boundary is represented by a solid line to show that the points on the line are included in the solution set.
To solve a linear inequality of the type above, one starts by drawing the boundary line. This boundary line partition the Cartesian coordinates system into two half-planes. One of them is the solution set. To determine which of the two half-planes is the solution set, one picks a point, called a **test point**, in one of the half-plane. If the chosen point is a solution to the linear inequality then the half-plane containing the point is the solution set. Otherwise, the half-plane not containing the point is the solution set.

Example 27.1
Solve graphically each of the following inequalities:
(a) $y \leq x - 2$
(b) $y < x - 2$
(c) $y \geq x - 2$
(d) $y > x - 2$.

Solution.
(a) First, we graph the line $y = x - 2$ as a solid line. The test point $(0, 0)$ does not satisfy the inequality so that the lower-half plane including the boundary line is the solution set. See Figure 27.1(a).
(b) The boundary line is a dashed-line. As in (a), the solution set is the lower-half plane. See Figure 27.1(b).
(c) The solution set is the upper-half plane together with the boundary line. See Figure 27.1(c).
(d) The solution set is the upper-half plane. See Figure 27.1(d) ∎

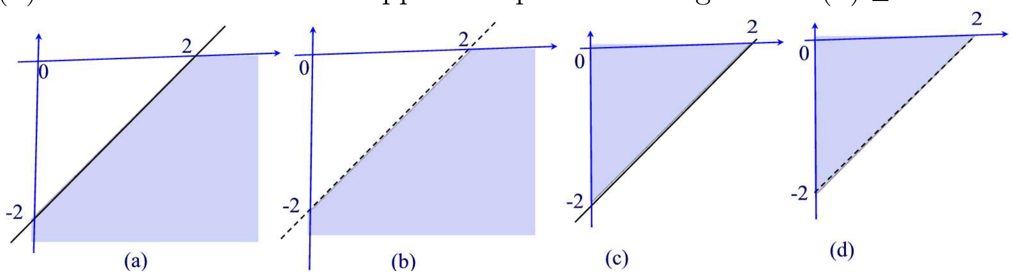


Figure 27.1

We next consider systems of linear inequalities. The solution set is the region in the Cartesian coordinate system consisting of all pairs that simultaneously satisfy all the inequalities in the system. The solution region is known as the **feasible region**.
To find the feasible region, we solve graphically each linear inequality in the system. The feasible region is the region where all the solution sets overlap. The intersection of two boundary lines is called a **corner point**.

Example 27.2
Solve the linear system

$$\left\{\begin{array}{c} -3x + 4y \leq 12 \\ x + 2y < 6 \\ -x + 5y \geq -5. \end{array}\right.$$

Solution.
We graph the solution set to each linear inequality on the same set of axes. The overlapping region is the triangle with corners $(0,3)$, $\left(-\frac{80}{11}, -\frac{27}{11}\right)$, $\left(\frac{40}{7}, \frac{1}{7}\right)$ as shown in Figure 27.2 ∎

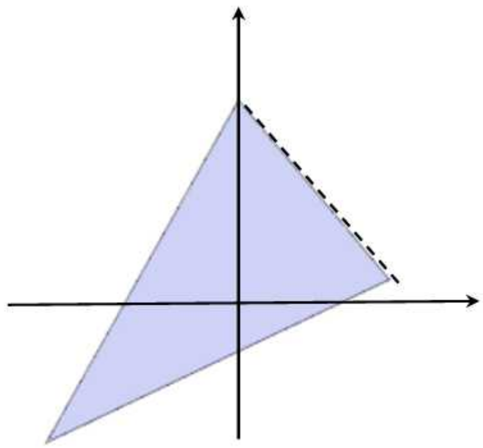


Figure 27.2

Practice Problems

Problem 27.1
Solve graphically $2x - 3y \leq 6$.

Problem 27.2
Solve graphically $x > 3$.

Problem 27.3
Solve graphically $y \leq 2$.

Problem 27.4
Solve graphically $2x + 5y > 20$.

Problem 27.5
Solve the system of inequalities:

$$\left\{\begin{array}{c} x - y < 1 \\ 2x + 3y \leq 12 \\ x \geq 0. \end{array}\right.$$

Problem 27.6
Solve the system of inequalities:

$$\left\{\begin{array}{c} x + 2y \leq 3 \\ -3x + y < 5 \\ -x + 8y \geq -23. \end{array}\right.$$

Problem 27.7
Solve the system of inequalities:

$$\left\{\begin{array}{c} y < 2x + 1 \\ x + 2y \geq -4. \end{array}\right.$$

Problem 27.8
Solve the system of inequalities:

$$\left\{\begin{array}{c} x > -2 \\ y \leq 4 \\ 3x + 4y \leq 24. \end{array}\right.$$

28 Improper Integrals

A very common mistake among students is when evaluating the integral $\int_{-1}^{1} \frac{1}{x} dx$. A non careful student will just argue as follows

$$\int_{-1}^{1} \frac{1}{x} dx = [\ln |x|]_{-1}^{1} = 0.$$

Unfortunately, that's not the right answer as we will see below. The important fact ignored here is that the integrand is not continuous at $x = 0$. Infact, $f(x) = \frac{1}{x}$ has an infinite discontinuity at $x = 0$.
Up to this point, the definite integral $\int_a^b f(x)dx$ only when (a) $f(x)$ is continuous on $[a, b]$,
(b) $[a, b]$ is of finite length.

Improper integrals are integrals in which one or both of these conditions are not met, i.e.,

(1) The interval of integration is infinite:

$$[a, \infty), (-\infty, b], (-\infty, \infty),$$

e.g.:

$$\int_1^\infty \frac{1}{x} dx.$$

(2) The integrand has an infinite discontinuity at some point c in the interval $[a, b]$, i.e. the integrand is unbounded near c :

$$\lim_{x \to c} f(x) = \pm\infty.$$

e.g.:

$$\int_0^1 \frac{1}{x} dx.$$

An improper integral is defined in terms of limits so it may exist or may not exist. If the limit exists, we say that the improper integral is **convergent.** Otherwise, the integral is **divergent**.
We will consider only improper integrals with positive integrands since they are the most common.

Infinite Intervals of Integration

The first type of improper integrals arises when the domain of integration is infinite but the integrand is still continuous in the domain of integration. In case one of the limits of integration is infinite, we define

$$\int_a^\infty f(x)dx = \lim_{b\to\infty} \int_a^b f(x)dx$$

or

$$\int_{-\infty}^b f(x)dx = \lim_{a\to-\infty} \int_a^b f(x)dx.$$

If both limits are infinite, we write

$$\int_{-\infty}^\infty f(x)dx = \lim_{R\to\infty} \int_{-R}^c f(x)dx + \lim_{R\to\infty} \int_c^R f(x)dx.$$

Example 28.1

Does the integral $\int_1^\infty \frac{1}{x^2}dx$ converge or diverge?

Solution.

We have

$$\int_1^\infty \frac{1}{x^2}dx = \lim_{b\to\infty} \int_1^b \frac{1}{x^2}dx = \lim_{b\to\infty} [-\frac{1}{x}]_1^b = \lim_{b\to\infty} (-\frac{1}{b} + 1) = 1.$$

In terms of area, the given integral represents the area under the graph of $f(x) = \frac{1}{x^2}$ from $x = 1$ and extending infinitely to the right. The above improper integral says the following. Let $b > 1$ and obtain the area shown in Figure 28.1.

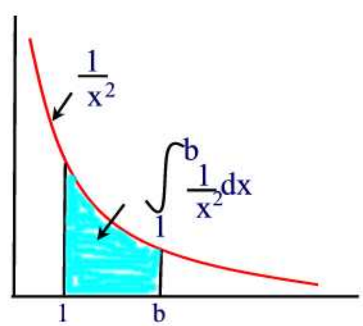


Figure 28.1

Then $\int_1^b \frac{1}{x^2}dx$ is the area under the graph of $f(x)$ from $x = 1$ to $x = b$. As b gets larger and larger this area is close to 1 ■

Example 28.2
Does the improper integral $\int_1^\infty \frac{1}{\sqrt{x}}dx$ converge or diverge?

Solution.
We have

$$\int_1^\infty \frac{1}{\sqrt{x}}dx = \lim_{b\to\infty} \int_1^b \frac{1}{\sqrt{x}}dx = \lim_{b\to\infty} [2\sqrt{x}]_1^b = \lim_{b\to\infty} (2\sqrt{b} - 2) = \infty.$$

So the improper integral is divergent ■

The following example generalizes the results of the previous two examples.

Example 28.3
Determine for which values of p the improper integral $\int_1^\infty \frac{1}{x^p}dx$ diverges.

Solution.
Suppose first that $p = 1$. Then

$$\begin{aligned}\int_1^\infty \frac{1}{x}dx &= \lim_{b\to\infty} \int_1^b \frac{1}{x}dx \\ &= \lim_{b\to\infty} [\ln|x|]_1^b = \lim_{b\to\infty} \ln b = \infty\end{aligned}$$

so the improper integral is divergent.
Now, suppose that $p \neq 1$. Then

$$\begin{aligned}\int_1^\infty \frac{1}{x^p}dx &= \lim_{b\to\infty} \int_1^b x^{-p}dx \\ &= \lim_{b\to\infty} [\frac{x^{-p+1}}{-p+1}]_1^b \\ &= \lim_{b\to\infty} (\frac{b^{-p+1}}{-p+1} - \frac{1}{-p+1}).\end{aligned}$$

If $p < 1$ then $-p+1 > 0$ so that $\lim_{b\to\infty} b^{-p+1} = \infty$ and therefore the improper integral is divergent. If $p > 1$ then $-p+1 < 0$ so that $\lim_{b\to\infty} b^{-p+1} = 0$ and hence the improper integral converges:

$$\int_1^\infty \frac{1}{x^p}dx = \frac{-1}{-p+1} \blacksquare$$

Example 28.4
For what values of c is the improper integral $\int_0^\infty e^{cx}dx$ convergent?

Solution.
We have

$$\begin{aligned}\int_0^\infty e^{cx}dx = &\lim_{b\to\infty}\int_0^b e^{cx}dx = \lim_{b\to\infty}\frac{1}{c}e^{cx}|_0^b\\ = &\lim_{b\to\infty}\frac{1}{c}(e^{cb}-1) = -\frac{1}{c}\end{aligned}$$

provided that $c < 0$. Otherwise, i.e. if $c \geq 0$, then the improper integral is divergent.$\blacksquare$

Remark 28.1
The previous two results are very useful and you may want to memorize them.

Example 28.5
Show that the improper integral $\int_{-\infty}^\infty \frac{1}{1+x^2}dx$ converges.

Solution.
We have

$$\begin{aligned}\int_{-\infty}^\infty \frac{1}{1+x^2}dx = &\lim_{a\to\infty}\int_{-a}^c \frac{1}{1+x^2}dx + \lim_{a\to\infty}\int_c^a \frac{1}{1+x^2}dx\\ = &\lim_{a\to\infty} \tan^{-1}x\Big]_{-a}^c + \lim_{a\to\infty}\tan^{-1}x\Big]_c^a\\ = &\lim_{a\to\infty}[\tan^{-1}a - \tan^{-1}-a]\\ = &\lim_{a\to\infty} 2\tan^{-1}a\\ = &2\frac{\pi}{2} = \pi \blacksquare\end{aligned}$$

Integrands with infinite discontinuity

Suppose $f(x)$ is unbounded at $x = a$, that is $\lim_{x \to a^+} f(x) = \infty$. Then we define

$$\int_a^b f(x)dx = \lim_{t \to a^+} \int_t^b f(x)dx.$$

Similarly, if $f(x)$ is unbounded at $x = b$, that is $\lim_{x \to b^-} f(x) = \infty$. Then we define

$$\int_a^b f(x)dx = \lim_{t \to b^-} \int_a^t f(x)dx.$$

Now, if $f(x)$ is unbounded at an interior point $a < c < b$ then we define

$$\int_a^b f(x)dx = \lim_{t \to c^-} \int_a^t f(x)dx + \lim_{t \to c^+} \int_t^b f(x)dx.$$

If both limits exist then the integral on the left-hand side converges. If one of the limits does not exist then we say that the improper integral is divergent.

Example 28.6

Show that the improper integral $\int_0^1 \frac{1}{\sqrt{x}} dx$ converges.

Solution.

Indeed,

$$\begin{aligned}\int_0^1 \frac{1}{\sqrt{x}} dx = &\lim_{t \to 0^+} \int_t^1 \frac{1}{\sqrt{x}} dx = \lim_{t \to 0^+} 2\sqrt{x}|_t^1 \\ = &\lim_{t \to 0^+} (2 - 2\sqrt{t}) = 2.\end{aligned}$$

In terms of area, we pick a $t > 0$ as shown in Figure 28.2. Then the shaded area is $\int_t^1 \frac{1}{\sqrt{x}} dx$. As t approaches 0 from the right, the area approaches the value 2 ■

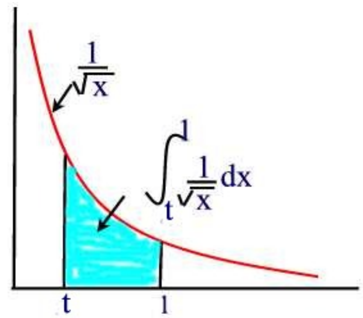


Figure 28.2

Example 28.7
Investigate the convergence of $\int_0^2 \frac{1}{(x-2)^2}dx$.

Solution.
We deal with this improper integral as follows

$$\int_0^2 \frac{1}{(x-2)^2}dx = \lim_{t\to 2^-}\int_0^t \frac{1}{(x-2)^2}dx = \lim_{t\to 2^-} -\frac{1}{(x-2)}|_0^t$$
$$= \lim_{t\to 2^-}(-\frac{1}{t-2} - \frac{1}{2}) = \infty.$$

So that the given improper integral is divergent ■

Example 28.8
Investigate the improper integral $\int_{-1}^1 \frac{1}{x}dx$.

Solution.
We first write
$$\int_{-1}^1 \frac{1}{x}dx = \int_{-1}^0 \frac{1}{x}dx + \int_0^1 \frac{1}{x}dx.$$
On one hand we have,

$$\int_{-1}^0 \frac{1}{x}dx = \lim_{t\to 0^-}\int_{-1}^t \frac{1}{x}dx = \lim_{t\to 0^-} \ln|x||_{-1}^t$$
$$= \lim_{t\to 0^-} \ln|t| = \infty.$$

This shows that the improper integral $\int_{-1}^{0}\frac{1}{x}dx$ is divergent and therefore the improper integral $\int_{-1}^{1}\frac{1}{x}dx$ is divergent ■

Improper Integrals of Mixed Type

Now, if the interval of integration is unbounded and the integrand is unbounded at one or more points inside the interval of integration we can evaluate the improper integral by decomposing it into a sum of an improper integral with finite interval but where the integrand is unbounded and an improper integral with an infinite interval. If each component integrals converges, then we say that the original integral converges to the sum of the values of the component integrals. If one of the component integrals diverges, we say that the entire integral diverges.

Example 28.9

Investigate the convergence of $\int_0^\infty \frac{1}{x^2}dx$.

Solution.

Note that the interval of integration is infinite and the function is undefined at $x = 0$. So we write

$$\int_0^\infty \frac{1}{x^2}dx = \int_0^1 \frac{1}{x^2}dx + \int_1^\infty \frac{1}{x^2}dx.$$

But

$$\int_0^1 \frac{1}{x^2}dx = \lim_{t\to 0^+}\int_t^1 \frac{1}{x^2}dx = \lim_{t\to 0^+} -\frac{1}{x}|_t^1 = \lim_{t\to 0^+}(\frac{1}{t}-1) = \infty.$$

Thus, $\int_0^1 \frac{1}{x^2}dx$ diverges and consequently the improper integral $\int_0^\infty \frac{1}{x^2}dx$ diverges ■

Comparison Tests for Improper Integrals

Sometimes it is difficult to find the exact value of an improper integral by antidifferentiation, for instance the integral $\int_0^\infty e^{-x^2}dx$. However, it is still possible to determine whether an improper integral converges or diverges. The idea is to compare the integral to one whose behavior we already know, such us

- the p-integral $\int_1^\infty \frac{1}{x^p}dx$ which converges for $p > 1$ and diverges otherwise;
- the integral $\int_0^\infty e^{cx}dx$ which converges for $c < 0$ and diverges for $c \geq 0$;
- the integral $\int_0^1 \frac{1}{x^p}dx$ which converges for $p < 1$ and diverges otherwise.

The comparison method consists of the following:

Theorem 28.1
Suppose that f and g are continuous and $0 \leq g(x) \leq f(x)$ for all $x \geq a$. Then

(a) if $\int_a^\infty f(x)dx$ is convergent, so is $\int_a^\infty g(x)dx$
(b) if $\int_a^\infty g(x)dx$ is divergent, so is $\int_a^\infty f(x)dx$.

This is only common sense: if the curve $y = g(x)$ lies below the curve $y = f(x)$, and the area of the region under the graph of $f(x)$ is finite, then of course so is the area of the region under the graph of $g(x)$. Similar results hold for the other types of improper integrals.

Example 28.10
Determine whether $\int_1^\infty \frac{1}{\sqrt{x^3+5}}dx$ converges.

Solution.
For $x \geq 1$ we have that $x^3 + 5 \geq x^3$ so that $\sqrt{x^3+5} \geq \sqrt{x^3}$. Thus, $\frac{1}{\sqrt{x^3+5}} \leq \frac{1}{\sqrt{x^3}}$. Letting $f(x) = \frac{1}{\sqrt{x^3}}$ and $g(x) = \frac{1}{\sqrt{x^3+5}}$ then we have that $0 \leq g(x) \leq f(x)$. From the previous section we know that $\int_1^\infty \frac{1}{x^{\frac{3}{2}}}dx$ is convergent, a p-integral with $p = \frac{3}{2} > 1$. By the comparison test, $\int_1^\infty \frac{1}{\sqrt{x^3+5}}dx$ is convergent ■

Example 28.11
Investigate the convergence of $\int_4^\infty \frac{dx}{\ln x - 1}$.

Solution.
For $x \geq 4$ we know that $\ln x - 1 < \ln x < x$. Thus, $\frac{1}{\ln x - 1} > \frac{1}{x}$. Let $g(x) = \frac{1}{x}$ and $f(x) = \frac{1}{\ln x - 1}$. Thus, $0 < g(x) \leq f(x)$. Since $\int_4^\infty \frac{1}{x}dx = \int_1^\infty \frac{1}{x}dx - \int_1^4 \frac{1}{x}dx$ and the integral $\int_1^\infty \frac{1}{x}dx$ is divergent being a p-integral with $p = 1$, the integral $\int_4^\infty \frac{1}{x}dx$ is divergent. By the comparison test $\int_4^\infty \frac{dx}{\ln x - 1}$ is divergent ■

Practice Problems

Problem 28.1
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_{-\infty}^{0} \frac{dx}{\sqrt{3-x}}.$$

Problem 28.2
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_{-1}^{1} \frac{e^x}{e^x - 1} dx.$$

Problem 28.3
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_{1}^{4} \frac{dx}{x-2}.$$

Problem 28.4
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_{1}^{10} \frac{dx}{\sqrt{10-x}}.$$

Problem 28.5
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_{-\infty}^{\infty} \frac{dx}{e^x + e^{-x}}.$$

Problem 28.6
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_{0}^{\infty} \frac{dx}{x^2+4}.$$

Problem 28.7
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_{-\infty}^{0} e^x dx.$$

Problem 28.8
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_0^{\infty} \frac{dx}{(x-5)^{\frac{1}{3}}}.$$

Problem 28.9
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_0^2 \frac{dx}{(x-1)^2}.$$

Problem 28.10
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_{-\infty}^{\infty} \frac{x}{x^2+9} dx.$$

Problem 28.11
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_0^1 \frac{4dx}{\sqrt{1-x^2}}.$$

Problem 28.12
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_0^{\infty} xe^{-x} dx.$$

Problem 28.13
Determine if the following integral is convergent or divergent. If it is convergent find its value.
$$\int_0^1 \frac{x^2}{\sqrt{1-x^3}} dx.$$

Problem 28.14
Determine if the following integral is convergent or divergent. If it is convergent find its value.

$$\int_1^2 \frac{x}{x-1} dx.$$

Problem 28.15
Investigate the convergence of $\int_4^\infty \frac{dx}{\ln x - 1}$.

Problem 28.16
Investigate the convergence of the improper integral $\int_1^\infty \frac{\sin x + 3}{\sqrt{x}} dx$.

Problem 28.17
Investigate the convergence of $\int_1^\infty e^{-\frac{1}{2}x^2} dx$.

29 Iterated Double Integrals

In this section, we see how to compute double integrals exactly using one-variable integrals.
Going back to the definition of the integral over a region as the limit of a double Riemann sum:

$$\begin{aligned}
\int_R f(x,y)dxdy &= \lim_{m,n\to\infty} \sum_{j=1}^m \sum_{i=1}^n f(x_i^*, y_j^*)\Delta x \Delta y \\
&= \lim_{m,n\to\infty} \sum_{j=1}^m \left(\sum_{i=1}^n f(x_i^*, y_j^*)\Delta x \right) \Delta y \\
&= \lim_{m,n\to\infty} \sum_{j=1}^m \Delta y \left(\sum_{i=1}^n f(x_i^*, y_j^*)\Delta x \right) \\
&= \lim_{m\to\infty} \sum_{j=1}^m \Delta y \int_a^b f(x, y_j^*)dx
\end{aligned}$$

We now let

$$F(y_j^*) = \int_a^b f(x, y_j^*)dx$$

and, substituting into the expression above, we obtain

$$\int_R f(x,y)dxdy = \lim_{m\to\infty} \sum_{j=1}^m F(y_j^*)\Delta y = \int_c^d F(y)dy = \int_c^d \int_a^b f(x,y)dxdy.$$

Thus, if f is continuous over a rectangle R then the integral of f over R can be expressed as an **iterated integral**. To evaluate this iterated integral, first perform the inside integral with respect to x, holding y constant, then integrate the result with respect to y.

Example 29.1
Compute $\int_0^{16} \int_0^8 \left(12 - \frac{x}{4} - \frac{y}{8}\right) dxdy$.

Solution.
We have

$$\begin{aligned}\int_0^{16}\int_0^8\left(12-\frac{x}{4}-\frac{y}{8}\right)dxdy=&\int_0^{16}\left(\int_0^8\left(12-\frac{x}{4}-\frac{y}{8}\right)dx\right)dy\\=&\int_0^{16}\left[12x-\frac{x^2}{8}-\frac{xy}{8}\right]_0^8dy\\=&\int_0^{16}(88-y)dy=88y-\left.\frac{y^2}{2}\right|_0^{16}=1280\ \blacksquare\end{aligned}$$

We note, that we can repeat the argument above for establishing the iterated integral, reversing the order of the summation so that we sum over j first and i second (i.e. integrate over y first and x second) so the result has the order of integration reversed. That is we can show that

$$\int_R f(x,y)dxdy=\int_a^b\int_c^d f(x,y)dydx.$$

Example 29.2
Compute $\int_0^8\int_0^{16}\left(12-\frac{x}{4}-\frac{y}{8}\right)dydx$.

Solution.
We have

$$\begin{aligned}\int_0^{8}\int_0^{16}\left(12-\frac{x}{4}-\frac{y}{8}\right)dydx=&\int_0^{8}\left(\int_0^{16}\left(12-\frac{x}{4}-\frac{y}{8}\right)dy\right)dx\\=&\int_0^{8}\left[12y-\frac{xy}{4}-\frac{y^2}{16}\right]_0^{16}dx\\=&\int_0^{8}(176-4x)dx=176x-2x^2\Big|_0^{8}=1280\ \blacksquare\end{aligned}$$

Iterated Integrals Over Non-Rectangular Regions
So far we looked at double integrals over rectangular regions. The problem with this is that most of the regions are not rectangular so we need to now look at the following double integral,

$$\int_R f(x,y)dxdy$$

where R is any region. We consider the two types of regions shown in Figure 29.1.

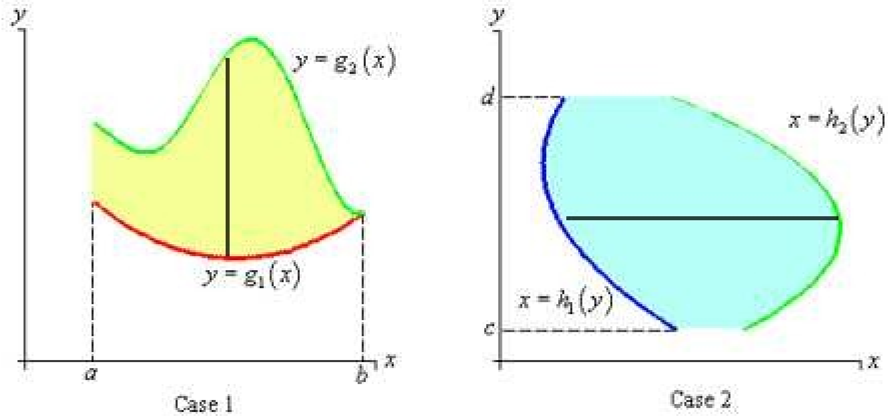


Figure 29.1

In Case 1, the iterated integral of f over R is defined by

$$\int_R f(x,y)dxdy = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)dydx$$

This means, that we are integrating using vertical strips from $g_1(x)$ to $g_2(x)$ and moving these strips from $x = a$ to $x = b$.
In case 2, we have

$$\int_R f(x,y)dxdy = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)dxdy$$

so we use horizontal strips from $h_1(y)$ to $h_2(y)$. Note that in both cases, the limits on the outer integral must always be constants.

Remark 29.1
Chosing the order of integration will depend on the problem and is usually determined by the function being integrated and the shape of the region R. The order of integration which results in the "simplest" evaluation of the integrals is the one that is preferred.

Example 29.3
Let $f(x,y) = xy$. Integrate $f(x,y)$ for the triangular region bounded by the $x-$axis, the $y-$axis, and the line $y = 2 - 2x$.

Solution.
Figure 29.2 shows the region of integration for this example.

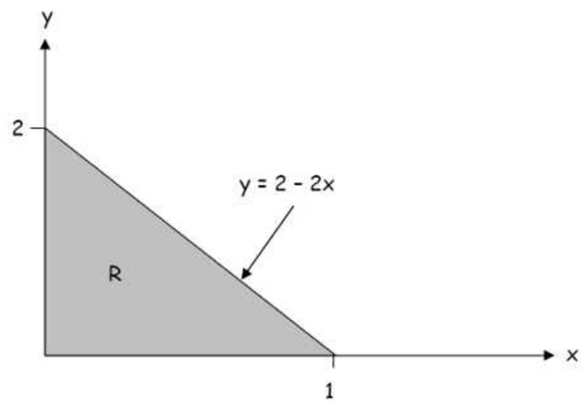


Figure 29.2

Graphically integrating over y first is equivalent to moving along the x axis from 0 to 1 and integrating from $y = 0$ to $y = 2 - 2x$. That is, summing up the vertical strips as shown in Figure 29.3(I).

$$
\begin{aligned}
\int_R xydxdy =& \int_0^1 \int_0^{2-2x} xydydx \\
=& \int_0^1 \left. \frac{xy^2}{2} \right|_0^{2-2x} dx = \frac{1}{2} \int_0^1 x(2-2x)^2 dx \\
=& 2\int_0^1 (x - 2x^2 + x^3)dx = 2\left[\frac{x^2}{2} - \frac{2}{3}x^3 + \frac{x^4}{4}\right]_0^1 = \frac{1}{6}
\end{aligned}
$$

If we choose to do the integral in the opposite order, then we need to invert the $y = 2 - 2x$ i.e. express x as function of y. In this case we get $x = 1 - \frac{1}{2}y$. Integrating in this order corresponds to integrating from $y = 0$ to $y = 2$ along horizontal strips ranging from $x = 0$ to $x = 1 - \frac{1}{2}y$, as shown in Figure

29.3(II)

$$\begin{aligned}
\int_R xydxdy =& \int_0^2 \int_0^{1-\frac{1}{2}y} xydxdy \\
=& \int_0^2 \left. \frac{x^2y}{2} \right|_0^{1-\frac{1}{2}y} dy = \frac{1}{2}\int_0^2 y(1-\frac{1}{2}y)^2 dy \\
=& \frac{1}{2}\int_0^2 (y - y^2 + \frac{y^3}{4})dy = \left. \frac{y^2}{4} - \frac{y^3}{6} + \frac{y^4}{32} \right|_0^2 = \frac{1}{6} \blacksquare
\end{aligned}$$

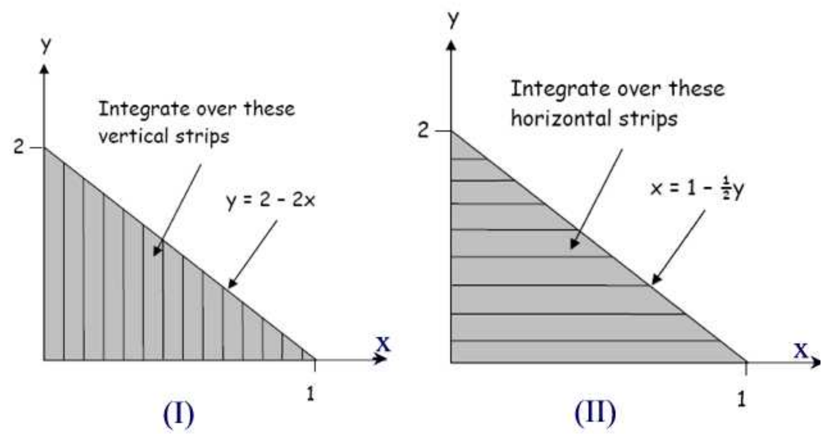


Figure 29.3

Example 29.4
Find $\int_R (4xy - y^3)dxdy$ where R is the region bounded by the curves $y = \sqrt{x}$ and $y = x^3$.

Solution.
A sketch of R is given in Figure 29.4. Using horizontal strips we can write

$$\begin{aligned}
\int_R (4xy - y^3)dxdy =& \int_0^1 \int_{y^2}^{\sqrt[3]{y}} (4xy - y^3)dxdy \\
=& \int_0^1 2x^2y - xy^3\Big|_{y^2}^{\sqrt[3]{y}} dy = \int_0^1 \left(2y^{\frac{5}{3}} - y^{\frac{10}{3}} - y^5\right) dy \\
=& \left. \frac{3}{4}y^{\frac{8}{3}} - \frac{3}{13}y^{\frac{13}{3}} - \frac{1}{6}y^6 \right|_0^1 = \frac{55}{156} \blacksquare
\end{aligned}$$

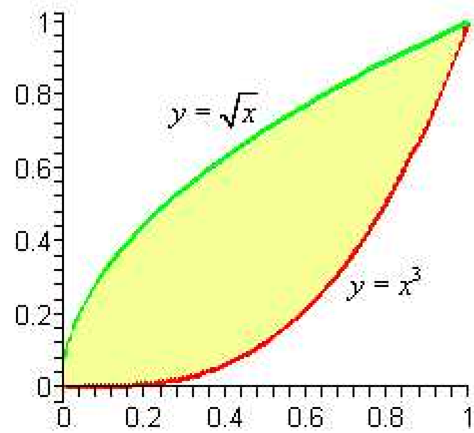


Figure 29.4

Example 29.5
Sketch the region of integration of $\int_0^2 \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} xydydx$

Solution.
A sketch of the region is given in Figure 29.5.

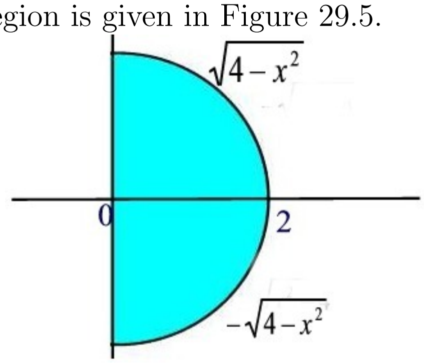


Figure 29.5

Practice Problems

Problem 29.1
Set up a double integral of $f(x, y)$ over the region given by $0 < x < 1; x < y < x + 1$.

Problem 29.2
Set up a double integral of $f(x, y)$ over the part of the unit square $0 \leq x \leq 1; 0 \leq y \leq 1$, on which $y \leq \frac{x}{2}$.

Problem 29.3
Set up a double integral of $f(x, y)$ over the part of the unit square on which both x and y are greater than 0.5.

Problem 29.4
Set up a double integral of $f(x, y)$ over the part of the unit square on which at least one of x and y is greater than 0.5.

Problem 29.5
Set up a double integral of $f(x, y)$ over the part of the region given by $0 < x < 50 - y < 50$ on which both x and y are greater than 20.

Problem 29.6
Set up a double integral of $f(x, y)$ over the set of all points (x, y) in the first quadrant with $|x - y| \leq 1$.

Problem 29.7
Evaluate $\int\int_R e^{-x-y} dxdy$, where R is the region in the first quadrant in which $x + y \leq 1$.

Problem 29.8
Evaluate $\int\int_R e^{-x-2y} dxdy$, where R is the region in the first quadrant in which $x \leq y$.

Problem 29.9
Evaluate $\int\int_R (x^2 + y^2) dxdy$, where R is the region $0 \leq x \leq y \leq L$.

Problem 29.10
Write as an iterated integral $\int\int_R f(x, y) dxdy$, where R is the region inside the unit square in which both coordinates x and y are greater than 0.5.

Problem 29.11
Evaluate $\int\int_R (x-y+1)dxdy$, where R is the region inside the unit square in which $x+y \geq 0.5$.

Problem 29.12
Evaluate $\int_0^1 \int_0^1 x\max(x,y)dydx$.

Continuous Random Variables

Continuous random variables are random quantities that are measured on a continuous scale. They can usually take on any value over some interval, which distinguishes them from discrete random variables, which can take on only a sequence of values, usually integers. Typically random variables that represent, for example, time or distance will be continuous rather than discrete.

30 Distribution Functions

We say that a random variable is **continuous** if there exists a nonnegative function f (not necessarily continuous) defined for all real numbers and having the property that for any set B of real numbers we have

$$\Pr(X \in B) = \int_B f(x)dx.$$

We call the function f the **probability density function** (abbreviated pdf) of the random variable X.
If we let $B = (-\infty, \infty)$ then

$$\int_{-\infty}^{\infty} f(x)dx = P[X \in (-\infty, \infty)] = 1.$$

Now, if we let $B = [a, b]$ then

$$\Pr(a \leq X \leq b) = \int_a^b f(x)dx.$$

That is, areas under the probability density function represent probabilities as illustrated in Figure 30.1.

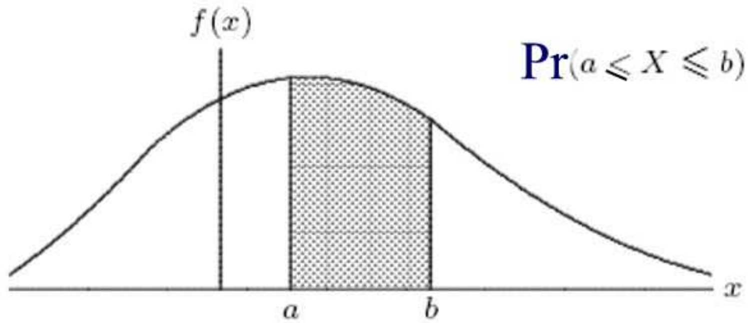


Figure 30.1

Now, if we let $a=b$ in the previous formula we find

$$\Pr(X=a)=\int_a^a f(x)dx=0.$$

It follows from this result that

$$\Pr(a\leq X<b)=\Pr(a<X\leq b)=\Pr(a<X<b)=\Pr(a\leq X\leq b).$$

and

$$\Pr(X\leq a)=\Pr(X<a)\quad\text{and}\quad\Pr(X\geq a)=\Pr(X>a).$$

The **cumulative distribution function** or simply the **distribution function** (abbreviated cdf) $F(t)$ of the random variable X is defined as follows

$$F(t)=\Pr(X\leq t)$$

i.e., $F(t)$ is equal to the probability that the variable X assumes values, which are less than or equal to t. From this definition we can write

$$F(t)=\int_{-\infty}^t f(y)dy.$$

Geometrically, $F(t)$ is the area under the graph of f to the left of t.

Example 30.1
Find the distribution functions corresponding to the following density functions:
(a) $f(x)=\frac{1}{\pi(1+x^2)},\quad -\infty<x<\infty$
(b) $f(x)=\frac{e^{-x}}{(1+e^{-x})^2},\quad -\infty<x<\infty$
(c) $f(x)=\frac{a-1}{(1+x)^a},\quad 0<x<\infty$
(d) $f(x)=k\alpha x^{\alpha-1}e^{-kx^\alpha},\quad 0<x<\infty, k>0, \alpha>0.$

Solution.
(a)

$$\begin{aligned}F(x) &= \int_{-\infty}^{x} \frac{1}{\pi(1+y^2)}dy\\ &= \left[\frac{1}{\pi}\arctan y\right]_{-\infty}^{x}\\ &=\frac{1}{\pi}\arctan x - \frac{1}{\pi}\cdot\frac{-\pi}{2}\\ &=\frac{1}{\pi}\arctan x + \frac{1}{2}\end{aligned}$$

(b)

$$\begin{aligned}F(x) &= \int_{-\infty}^{x} \frac{e^{-y}}{(1+e^{-y})^2}dy\\ &= \left[\frac{1}{1+e^{-y}}\right]_{-\infty}^{x}\\ &=\frac{1}{1+e^{-x}}\end{aligned}$$

(c) For $x \geq 0$

$$\begin{aligned}F(x) &= \int_{-\infty}^{x} \frac{a-1}{(1+y)^a}dy\\ &= \left[-\frac{1}{(1+y)^{a-1}}\right]_{0}^{x}\\ &=1 - \frac{1}{(1+x)^{a-1}}\end{aligned}$$

For $x < 0$ it is obvious that $F(x) = 0$, so we could write the result in full as

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - \frac{1}{(1+x)^{a-1}} & x \geq 0. \end{cases}$$

(d) For $x \geq 0$

$$\begin{aligned}F(x) &= \int_{0}^{x} k\alpha y^{\alpha-1}e^{-ky^{\alpha}}dy\\ &= \left[-e^{-ky^{\alpha}}\right]_{0}^{x}\\ &=1 - e^{-kx^{\alpha}}\end{aligned}$$

For $x < 0$ we have $F(x) = 0$ so that

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - ke^{-kx^\alpha} & x \geq 0 \end{cases} \blacksquare$$

Next, we list the properties of the cumulative distribution function $F(t)$ for a continuous random variable X.

Theorem 30.1
The cumulative distribution function of a continuous random variable X satisfies the following properties:
(a) $0 \leq F(t) \leq 1$.
(b) $F(t)$ is a non-decreasing function, i.e. if $a < b$ then $F(a) \leq F(b)$.
(c) $F(t) \to 0$ as $t \to -\infty$ and $F(t) \to 1$ as $t \to \infty$.
(d) $\Pr(a < X \leq b) = F(b) - F(a)$.
(e) F is continuous.
(f) $F'(t) = f(t)$ whenever the derivative exists.

Proof.
Properties $(a) - (d)$ were established in Section 25. For part (e), we know that F is right continuous (See Proposition 25.3). Left-continuity follows from Example 25.4 and the fact that $\Pr(X = a) = 0$. Part (f) is the result of applying the Second Fundamental Theorem of Calculus to the function $F(x) = \int_{-\infty}^{x} f(t)dt$ $\blacksquare$

Figure 30.2 illustrates a representative cdf.

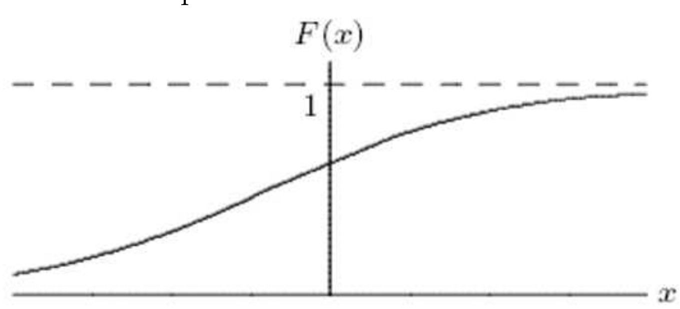


Figure 30.2

Remark 30.1
It is important to keep in mind that a pdf does not represent a probability.

However, it can be used as a measure of how likely it is that the random variable will be near a. To see this, let $\epsilon > 0$ be a small positive number. Then

$$\Pr(a \leq X \leq a+\epsilon) = F(a+\epsilon) - F(a) = \int_a^{a+\epsilon} f(t)dt \approx \epsilon f(a)$$

In particular,

$$\Pr\left(a - \frac{\epsilon}{2} \leq X \leq a + \frac{\epsilon}{2}\right) = \epsilon f(a).$$

This means that the probability that X will be contained in an interval of length ϵ around the point a is approximately $\epsilon f(a)$.

Remark 30.2
By Theorem 30.1 (c) and (f), $\lim_{t\to-\infty} f(t) = 0 = \lim_{t\to\infty} f(t)$. This follows from the fact that the graph of $F(t)$ levels off when $t \to \pm\infty$. That is, $\lim_{t\to\pm\infty} F'(t) = 0$.

Example 30.2
Suppose that the function $f(t)$ defined below is the density function of some random variable X.

$$f(t) = \begin{cases} e^{-t} & t \geq 0, \\ 0 & t < 0. \end{cases}$$

Compute $\Pr(-10 \leq X \leq 10)$.

Solution.

$$\begin{aligned}
\Pr(-10 \leq X \leq 10) &= \int_{-10}^{10} f(t)dt \\
&= \int_{-10}^{0} f(t)dt + \int_0^{10} f(t)dt \\
&= \int_0^{10} e^{-t}dt \\
&= -e^{-t}\Big|_0^{10} = 1 - e^{-10} \blacksquare
\end{aligned}$$

A pdf need not be continuous, as the following example illustrates.

Example 30.3
(a) Determine the value of c so that the following function is a pdf.

$$f(x) = \begin{cases} \frac{15}{64} + \frac{x}{64} & -2 \le x \le 0 \\ \frac{3}{8} + cx & 0 < x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

(b) Determine $\Pr(-1 \le X \le 1)$.
(c) Find $F(x)$.

Solution.
(a) Observe that f is discontinuous at the points -2 and 0, and is potentially also discontinuous at the point 3. We first find the value of c that makes f a pdf.

$$\begin{aligned} 1 =& \int_{-2}^{0} \left(\frac{15}{64} + \frac{x}{64} \right) dx + \int_{0}^{3} \left(\frac{3}{8} + cx \right) dx \\ =& \left[\frac{15}{64}x + \frac{x^2}{128} \right]_{-2}^{0} + \left[\frac{3}{8}x + \frac{cx^2}{2} \right]_{0}^{3} \\ =& \frac{30}{64} - \frac{2}{64} + \frac{9}{8} + \frac{9c}{2} \\ =& \frac{100}{64} + \frac{9c}{2} \end{aligned}$$

Solving for c we find $c = -\frac{1}{8}$.
(b) The probability $\Pr(-1 \le X \le 1)$ is calculated as follows.

$$\Pr(-1 \le X \le 1) = \int_{-1}^{0} \left(\frac{15}{64} + \frac{x}{64} \right) dx + \int_{0}^{1} \left(\frac{3}{8} - \frac{x}{8} \right) dx = \frac{69}{128}$$

(c) For $-2 \le x \le 0$ we have

$$F(x) = \int_{-2}^{x} \left(\frac{15}{64} + \frac{t}{64} \right) dt = \frac{x^2}{128} + \frac{15}{64}x + \frac{7}{16}$$

and for $0 < x \le 3$

$$F(x) = \int_{-2}^{0} \left(\frac{15}{64} + \frac{x}{64} \right) dx + \int_{0}^{x} \left(\frac{3}{8} - \frac{t}{8} \right) dt = \frac{7}{16} + \frac{3}{8}x - \frac{1}{16}x^2.$$

Hence the full cdf is

$$F(x) = \begin{cases} 0 & x < -2 \\ \frac{x^2}{128} + \frac{15}{64}x + \frac{7}{16} & -2 \le x \le 0 \\ \frac{7}{16} + \frac{3}{8}x - \frac{1}{16}x^2 & 0 < x \le 3 \\ 1 & x > 3 \end{cases}$$

Observe that at all points of discontinuity of the pdf, the cdf is continuous. That is, even when the pdf is discontinuous, the cdf is continuous ∎

Practice Problems

Problem 30.1
Determine the value of c so that the following function is a pdf.

$$f(x) = \begin{cases} \frac{c}{(x+1)^3} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Problem 30.2
Let X denote the length of time (in minutes) of using a computer at a public library with pdf given by

$$f(x) = \begin{cases} \frac{1}{5}e^{-\frac{x}{5}} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(a) What is the probability of using a computer for more than 10 minutes.
(b) Find the probability of using a computer between 5 and 10 minutes.
(c) Find the cumulative distribution function of X.

Problem 30.3
A probability student is always late to class and arrives within ten minutes after the start of the class. Let X be the time that elapses between the start of the class and the time the student arrives to class with a probability density function

$$f(x) = \begin{cases} kx^2 & 0 \leq x \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

where $k > 0$ is a constant. Compute the value of k and then find the probability that the student arrives less than 3 minutes after the start of the class.

Problem 30.4
The lifetime X of a battery (in hours) has a density function given by

$$f(x) = \begin{cases} 2x & 0 \leq x < \frac{1}{2} \\ \frac{3}{4} & 2 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

Find the probability that a battery will last for more than 15 minutes?

Problem 30.5
Let $F : \mathbb{R} \to \mathbb{R}$ be a function defined by

$$F(x) = \begin{cases} 0 & x < 0 \\ x/2 & 0 \leq x < 1 \\ (x+2)/6 & 1 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

(a) Show that F satisfies conditions (a),(b),(c), and (e) of Theorem 30.1.
(b) Find the probability density function $f(x)$.

Problem 30.6
The amount of time X (in minutes) it takes a person standing in line at a post office to reach the counter is described by the continuous probability function:

$$f(x) = \begin{cases} kxe^{-x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.
(a) Dtermine the value of k.
(b) What is the probability that a person takes more than 1 minute to reach the counter?

Problem 30.7
A random variable X has the cumulative distribution function

$$F(x) = \frac{e^x}{e^x + 1}.$$

(a) Find the probability density function.
(b) Find $\Pr(0 \leq X \leq 1)$.

Problem 30.8
A commercial water distributor supplies an office with gallons of water once a week. Suppose that the weekly supplies in tens of gallons is a random variable with pdf

$$f(x) = \begin{cases} 5(1-x)^4 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Approximately, how many gallons should be delivered in one week so that the probability of the supply is 0.1?

Problem 30.9 ‡
The loss due to a fire in a commercial building is modeled by a random variable X with density function

$$f(x) = \begin{cases} 0.005(20-x) & 0 < x < 20 \\ 0 & \text{otherwise} \end{cases}$$

Given that a fire loss exceeds 8, what is the probability that it exceeds 16 ?

Problem 30.10 ‡
The lifetime of a machine part has a continuous distribution on the interval $(0, 40)$ with probability density function f, where $f(x)$ is proportional to $(10+x)^{-2}$.
Calculate the probability that the lifetime of the machine part is less than 6.

Problem 30.11 ‡
A group insurance policy covers the medical claims of the employees of a small company. The value, V, of the claims made in one year is described by

$$V = 100000Y$$

where Y is a random variable with density function

$$f(x) = \begin{cases} k(1-y)^4 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.
What is the conditional probability that V exceeds 40,000, given that V exceeds 10,000?

Problem 30.12 ‡
An insurance company insures a large number of homes. The insured value, X, of a randomly selected home is assumed to follow a distribution with density function

$$f(x) = \begin{cases} 3x^{-4} & x > 1 \\ 0 & \text{otherwise} \end{cases}$$

Given that a randomly selected home is insured for at least 1.5, what is the probability that it is insured for less than 2 ?

Problem 30.13 ‡

An insurance policy pays for a random loss X subject to a deductible of C, where $0 < C < 1$. The loss amount is modeled as a continuous random variable with density function

$$f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Given a random loss X, the probability that the insurance payment is less than 0.5 is equal to 0.64 . Calculate C.

Problem 30.14

Let X_1, X_2, X_3 be three independent, identically distributed random variables each with density function

$$f(x) = \begin{cases} 3x^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Let $Y = \max\{X_1, X_2, X_3\}$. Find $\Pr(Y > \frac{1}{2})$.

Problem 30.15

Let X have the density function

$$f(x) = \begin{cases} \frac{3x^2}{\theta^3} & 0 < x < \theta \\ 0\text{otherwise} & \end{cases}$$

If $\Pr(X > 1) = \frac{7}{8}$, find the value of θ.

31 Expectation and Variance

As with discrete random variables, the expected value of a continuous random variable is a measure of location. It defines the balancing point of the distribution.
Suppose that a continuous random variable X has a density function $f(x)$ defined in $[a,b]$. Let's try to estimate $E(X)$ by cutting $[a,b]$ into n equal subintervals, each of width Δx, so $\Delta x = \frac{(b-a)}{n}$. Let $x_i = a + i\Delta x, i = 0, 1, ..., n,$ be the partition points between the subintervals. Then, the probability of X assuming a value in $[x_i, x_{i+1}]$ is

$$\Pr(x_i \le X \le x_{i+1}) = \int_{x_i}^{x_{i+1}} f(x)dx \approx \Delta x f\left(\frac{x_i + x_{i+1}}{2}\right)$$

where we used the midpoint rule to estimate the integral. An estimate of the desired expectation is approximately

$$E(X) \approx \sum_{i=0}^{n-1} \left(\frac{x_i + x_{i+1}}{2}\right) \Delta x f\left(\frac{x_i + x_{i+1}}{2}\right).$$

A better estimate is obtained by letting $n \to \infty$. Thus, we obtain

$$E(X) = \int_a^b x f(x)dx.$$

The above argument applies if either a or b are infinite. In this case, one has to make sure that all improper integrals in question converge.
Since the domain of f consists of all real numbers, we define the **expected value** of X by the improper integral

$$E(X) = \int_{-\infty}^{\infty} x f(x)dx$$

provided that the improper integral converges.

Example 31.1
Find $E(X)$ when the density function of X is

$$f(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Solution.
Using the formula for $E(X)$ we find

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^1 2x^2 dx = \frac{2}{3} \blacksquare$$

Example 31.2
A continuous random variable has the pdf

$$f(x) = \begin{cases} \frac{600}{x^2}, & 100 < x < 120 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Determine the mean and variance of X.
(c) Find $\Pr(X > 110)$.

Solution.
(a) We have

$$E(X) = \int_{100}^{120} x \cdot 600x^{-2} dx = 600 \ln x|_{100}^{120} \approx 109.39$$

and

$$\sigma_X^2 = E(X^2) - (E(X))^2 = \int_{100}^{120} x^2 \cdot 600x^{-2} dx - 109.39^2 \approx 33.19.$$

(b) The desired probability is

$$\Pr(X > 110) = \int_{110}^{120} 600x^{-2} dx = \frac{5}{11} \blacksquare$$

Sometimes for technical purposes the following theorem is useful. It expresses the expectation in terms of an integral of probabilities. It is most often used for random variables X that have only positive values; in that case the second term is of course zero.

Theorem 31.1
Let X be a continuous random variable with probability density function f. Then

$$E(X) = \int_0^{\infty} \Pr(X > y) dy - \int_{-\infty}^{0} \Pr(X < y) dy.$$

Proof.
From the definition of $E(X)$ we have

$$\begin{aligned}E(X) =& \int_0^\infty xf(x)dx + \int_{-\infty}^0 xf(x)dx\\ =& \int_0^\infty \int_{y=0}^{y=x} dyf(x)dx - \int_{-\infty}^0 \int_{y=x}^{y=0} dyf(x)dx\end{aligned}$$

Interchanging the order of integration as shown in Figure 31.1 we can write

$$\int_0^\infty \int_{y=0}^{y=x} dyf(x)dx = \int_0^\infty \int_y^\infty f(x)dxdy$$

and

$$\int_{-\infty}^0 \int_{y=x}^{y=0} dyf(x)dx = \int_{-\infty}^0 \int_{-\infty}^y f(x)dxdy.$$

The result follows by putting the last two equations together and recalling that

$$\int_y^\infty f(x)dx = \Pr(X > y) \quad \text{and} \quad \int_{-\infty}^y f(x)dx = \Pr(X < y) \blacksquare$$

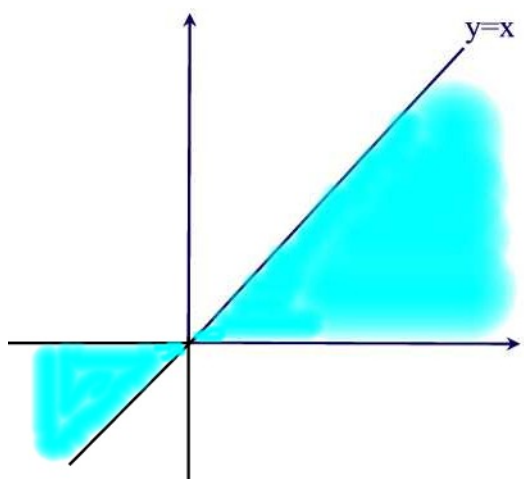


Figure 31.1

Note that if X is a continuous random variable and g is a function defined for the values of X and with real values, then $Y = g(X)$ is also a random variable. The following theorem is particularly important and convenient. If a random variable $Y = g(X)$ is expressed in terms of a continuous random variable, then this theorem gives the expectation of Y in terms of probabilities associated to X.

Theorem 31.2
If X is continuous random variable with a probability density function $f(x)$, and if $Y = g(X)$ is a function of the random variable, then the expected value of the function $g(X)$ is

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

Proof.
By the previous theorem we have

$$E(g(X)) = \int_0^{\infty} P[g(X) > y]dy - \int_{-\infty}^{0} P[g(X) < y]dy.$$

If we let $B_y = \{x : g(x) > y\}$ then from the definition of a continuous random variable we can write

$$P[g(X) > y] = \int_{B_y} f(x)dx = \int_{\{x:g(x)>y\}} f(x)dx.$$

Thus,

$$E(g(X)) = \int_0^{\infty} \left[\int_{\{x:g(x)>y\}} f(x)dx\right] dy - \int_{-\infty}^{0} \left[\int_{\{x:g(x)<y\}} f(x)dx\right] dy.$$

Now we can interchange the order of integration to obtain

$$\begin{aligned} E(g(X)) &= \int_{\{x:g(x)>0\}} \int_0^{g(x)} f(x)dydx - \int_{\{x:g(x)<0\}} \int_{g(x)}^{0} f(x)dydx \\ &= \int_{\{x:g(x)>0\}} g(x)f(x)dx + \int_{\{x:g(x)<0\}} g(x)f(x)dx = \int_{-\infty}^{\infty} g(x)f(x)dx. \end{aligned}$$

Figure 31.2 helps understanding the process of interchanging the order of integration that we used in the proof above ■

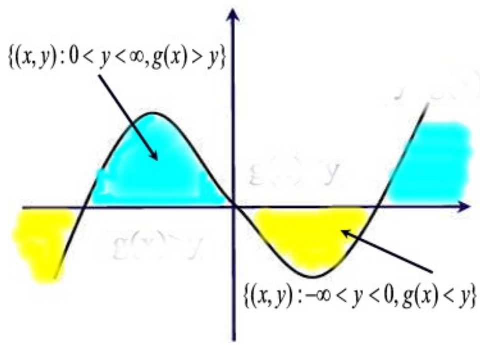


Figure 31.2

Example 31.3
Let T be a continuous random variable with pdf

$$f(t) = \begin{cases} \frac{1}{10}e^{-\frac{t}{10}}, & t \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Define the continuous random variable by

$$X = \begin{cases} 100 & 0 < T \leq 1 \\ 50 & 1 < T \leq 3 \\ 0 & T > 3. \end{cases}$$

Find $E(X)$.

Solution.
By Theorem 31.2, we have

$$\begin{aligned} E(X) =& \int_0^1 100\frac{1}{10}e^{-\frac{t}{10}}dt + \int_1^3 50\frac{1}{10}e^{-\frac{t}{10}}dt \\ =& 100(1 - e^{-\frac{1}{10}}) + 50(e^{-\frac{1}{10}} - e^{-\frac{3}{10}}) \\ =& 100 - 50e^{-\frac{1}{10}} - 50e^{-\frac{3}{10}} \blacksquare \end{aligned}$$

Example 31.4 ‡
An insurance policy reimburses a loss up to a benefit limit of 10 . The policyholder's loss, X, follows a distribution with density function:

$$f(x) = \begin{cases} \frac{2}{x^3} & x > 1 \\ 0 & \text{otherwise.} \end{cases}$$

What is the expected value of the benefit paid under the insurance policy?

Solution.
Let Y denote the claim payments. Then

$$Y = \begin{cases} X & 1 < X \leq 10 \\ 10 & X \geq 10 \end{cases}$$

It follows that

$$\begin{aligned} E(Y) &= \int_1^{10} x\frac{2}{x^3}dx + \int_{10}^{\infty} 10\frac{2}{x^3}dx \\ &= -\frac{2}{x}\Big|_1^{10} - \frac{10}{x^2}\Big|_{10}^{\infty} = 1.9 \blacksquare \end{aligned}$$

As a first application of Theorem 31.2, we have

Corollary 31.1
For any constants a and b

$$E(aX+b) = aE(X) + b.$$

Proof.
Let $g(x) = ax + b$ in Theorem 31.2 to obtain

$$\begin{aligned} E(aX+b) &= \int_{-\infty}^{\infty} (ax+b)f(x)dx \\ &= a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx \\ &= aE(X) + b \blacksquare \end{aligned}$$

Example 31.5 ‡
Claim amounts for wind damage to insured homes are independent random variables with common density function

$$f(x) = \begin{cases} \frac{3}{x^4} & x > 1 \\ 0 & \text{otherwise} \end{cases}$$

where x is the amount of a claim in thousands.
Suppose 3 such claims will be made, what is the expected value of the largest of the three claims?

Solution.
Note for any of the random variables the cdf is given by

$$F(x) = \int_1^x \frac{3}{t^4}dt = 1 - \frac{1}{x^3}, \quad x > 1$$

Next, let X_1, X_2, and X_3 denote the three claims made that have this distribution. Then if Y denotes the largest of these three claims, it follows that the cdf of Y is given by

$$\begin{aligned} F_Y(y) =& P[(X_1 \le y) \cap (X_2 \le y) \cap (X_3 \le y)] \\ =& \Pr(X_1 \le y)\Pr(X_2 \le y)\Pr(X_3 \le y) \\ =& \left(1 - \frac{1}{y^3}\right)^3, \quad y > 1 \end{aligned}$$

The pdf of Y is obtained by differentiating $F_Y(y)$

$$f_Y(y) = 3\left(1 - \frac{1}{y^3}\right)^2 \left(\frac{3}{y^4}\right) = \left(\frac{9}{y^4}\right)\left(1 - \frac{1}{y^3}\right)^2.$$

Finally,

$$\begin{aligned} E(Y) =& \int_1^\infty \left(\frac{9}{y^3}\right)\left(1 - \frac{1}{y^3}\right)^2 dy = \int_1^\infty \left(\frac{9}{y^3}\right)\left(1 - \frac{2}{y^3} + \frac{1}{y^6}\right) dy \\ =& \int_1^\infty \left(\frac{9}{y^3} - \frac{18}{y^6} + \frac{9}{y^9}\right) dy = \left[-\frac{9}{2y^2} + \frac{18}{5y^5} - \frac{9}{8y^8}\right]_1^\infty \\ =& 9\left[\frac{1}{2} - \frac{2}{5} + \frac{1}{8}\right] \approx 2.025 \text{ (in thousands)} \blacksquare \end{aligned}$$

Example 31.6 ‡
A manufacturer's annual losses follow a distribution with density function

$$f(x) = \begin{cases} \frac{2.5(0.6)^{2.5}}{x^{3.5}} & x > 0.6 \\ 0 & \text{otherwise.} \end{cases}$$

To cover its losses, the manufacturer purchases an insurance policy with an annual deductible of 2.
What is the mean of the manufacturer's annual losses not paid by the insurance policy?

Solution.
Let Y denote the manufacturer's retained annual losses. Then

$$Y = \begin{cases} X & 0.6 < X \leq 2 \\ 2 & X > 2 \end{cases}$$

Therefore,

$$\begin{aligned}
E(Y) =& \int_{0.6}^{2} x\left[\frac{2.5(0.6)^{2.5}}{x^{3.5}}\right]dx + \int_{2}^{\infty} 2\left[\frac{2.5(0.6)^{2.5}}{x^{3.5}}\right]dx \\
=& \int_{0.6}^{2} \left[\frac{2.5(0.6)^{2.5}}{x^{2.5}}\right]dx - \left.\frac{2(0.6)^{2.5}}{x^{2.5}}\right|_{2}^{\infty} \\
=& -\left.\frac{2.5(0.6)^{2.5}}{1.5x^{1.5}}\right|_{0.6}^{2} + \frac{2(0.6)^{2.5}}{2^{2.5}} \\
=& -\frac{2.5(0.6)^{2.5}}{1.5(2)^{1.5}} + \frac{2.5(0.6)^{2.5}}{1.5(0.6)^{1.5}} + \frac{2(0.6)^{2.5}}{2^{2.5}} \approx 0.9343 \blacksquare
\end{aligned}$$

The **variance** of a random variable is a measure of the "spread" of the random variable about its expected value. In essence, it tells us how much variation there is in the values of the random variable from its mean value. The variance of the random variable X, is determined by calculating the expectation of the function $g(X) = (X - E(X))^2$. That is,

$$\text{Var}(X) = E\left[(X - E(X))^2\right].$$

Theorem 31.3
(a) An alternative formula for the variance is given by

$$\text{Var}(X) = E(X^2) - [E(X)]^2.$$

(b) For any constants a and b, $\text{Var}(aX + b) = a^2\text{Var}(X)$.

Proof.
(a) By Theorem 31.2 we have

$$\begin{aligned}
\text{Var}(X) =& \int_{-\infty}^{\infty} (x - E(X))^2 f(x)dx \\
=& \int_{-\infty}^{\infty} (x^2 - 2xE(X) + (E(X))^2)f(x)dx \\
=& \int_{-\infty}^{\infty} x^2 f(x)dx - 2E(X)\int_{-\infty}^{\infty} xf(x)dx + (E(X))^2\int_{-\infty}^{\infty} f(x)dx \\
=& E(X^2) - (E(X))^2
\end{aligned}$$

(b) We have

$$\mathrm{Var}(aX+b) = E[(aX+b-E(aX+b))^2] = E[a^2(X-E(X))^2] = a^2\mathrm{Var}(X) \blacksquare$$

Example 31.7
Let X be a random variable with probability density function

$$f(x) = \begin{cases} 2-4|x| & -\frac{1}{2} < x < \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the variance of X.
(b) Find the c.d.f. $F(x)$ of X.

Solution.
(a) Since the function $xf(x)$ is odd in $-\frac{1}{2} < x < \frac{1}{2}$, we have $E(X) = 0$. Thus,

$$\begin{aligned}\mathrm{Var}(X) =& E(X^2) = \int_{-\frac{1}{2}}^{0} x^2(2+4x)dx + \int_0^{\frac{1}{2}} x^2(2-4x)dx \\ =& \frac{1}{24}.\end{aligned}$$

(b) Since the range of f is the interval $(-\frac{1}{2}, \frac{1}{2})$, we have $F(x) = 0$ for $x \le -\frac{1}{2}$ and $F(x) = 1$ for $x \ge \frac{1}{2}$. Thus it remains to consider the case when $-\frac{1}{2} < x < \frac{1}{2}$. For $-\frac{1}{2} < x \le 0$,

$$F(x) = \int_{-\frac{1}{2}}^{x} (2+4t)dt = 2x^2 + 2x + \frac{1}{2}.$$

For $0 \le x < \frac{1}{2}$, we have

$$F(x) = \int_{-\frac{1}{2}}^{0} (2+4t)dt + \int_0^x (2-4t)dt = -2x^2 + 2x + \frac{1}{2}.$$

Combining these cases, we get

$$F(x) = \begin{cases} 0 & x < -\frac{1}{2} \\ 2x^2 + 2x + \frac{1}{2} & -\frac{1}{2} \le x < 0 \\ -2x^2 + 2x + \frac{1}{2} & 0 \le x < \frac{1}{2} \\ 1 & x \ge \frac{1}{2} \blacksquare \end{cases}$$

Example 31.8
Let X be a continuous random variable with pdf

$$f(x) = \begin{cases} 4xe^{-2x}, & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

For this example, you might find the identity $\int_0^\infty t^n e^{-t} dt = n!$ useful.
(a) Find $E(X)$.
(b) Find the variance of X.
(c) Find the probability that $X < 1$.

Solution.
(a) Using the substitution $t = 2x$ we find

$$E(X) = \int_0^\infty 4x^2 e^{-2x} dx = \frac{1}{2}\int_0^\infty t^2 e^{-t} dt = \frac{2!}{2} = 1.$$

(b) First, we find $E(X^2)$. Again, letting $t = 2x$ we find

$$E(X^2) = \int_0^\infty 4x^3 e^{-2x} dx = \frac{1}{4}\int_0^\infty t^3 e^{-t} dt = \frac{3!}{4} = \frac{3}{2}.$$

Hence,

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{3}{2} - 1 = \frac{1}{2}.$$

(c) We have

$$\begin{aligned} \Pr(X < 1) =& \Pr(X \le 1) = \int_0^1 4xe^{-2x} dx = \int_0^2 te^{-t} dt \\ =& -(t+1)e^{-t}\Big|_0^2 = 1 - 3e^{-2} \blacksquare \end{aligned}$$

As in the case of discrete random variable, it is easy to establish the formula

$$\text{Var}(aX) = a^2 \text{Var}(X)$$

Example 31.9
Let X be the random variable representing the cost of maintaining a car. Suppose that $E(X) = 200$ and $\text{Var}(X) = 260$. If a tax of 20% is introduced on all items associated with the maintenance of the car, what will the variance of the cost of maintaining a car be?

Solution.
The new cost is $1.2X$, so its variance is $\text{Var}(1.2X) = 1.2^2\text{Var}(X) = (1.44)(260) = 374$. ■

Finally, we define the **standard deviation** X to be the square root of the variance.

Example 31.10
A random variable has a **Pareto** distribution with parameters $\alpha > 0$ and $x_0 > 0$ if its density function has the form

$$f(x) = \begin{cases} \frac{\alpha x_0^\alpha}{x^{\alpha+1}} & x > x_0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that $f(x)$ is indeed a density function.
(b) Find $E(X)$ and $\text{Var}(X)$.

Solution.
(a) By definition $f(x) > 0$. Also,

$$\int_{x_0}^{\infty} f(x)dx = \int_{x_0}^{\infty} \frac{\alpha x_0^\alpha}{x^{\alpha+1}} dx = -\left(\frac{x_0}{x}\right)\Big|_{x_0}^{\infty} = 1$$

(b) We have

$$E(X) = \int_{x_0}^{\infty} xf(x)dx = \int_{x_0}^{\infty} \frac{\alpha x_0^\alpha}{x^\alpha} dx = \frac{\alpha}{1-\alpha}\left(\frac{x_0^\alpha}{x^{\alpha-1}}\right)\Big|_{x_0}^{\infty} = \frac{\alpha x_0}{\alpha - 1}$$

provided $\alpha > 1$. Similarly,

$$E(X^2) = \int_{x_0}^{\infty} x^2 f(x)dx = \int_{x_0}^{\infty} \frac{\alpha x_0^\alpha}{x^{\alpha-1}} dx = \frac{\alpha}{2-\alpha}\left(\frac{x_0^\alpha}{x^{\alpha-2}}\right)\Big|_{x_0}^{\infty} = \frac{\alpha x_0^2}{\alpha - 2}$$

provided $\alpha > 2$. Hence,

$$\text{Var}(X) = \frac{\alpha x_0^2}{\alpha - 2} - \frac{\alpha^2 x_0^2}{(\alpha-1)^2} = \frac{\alpha x_0^2}{(\alpha-2)(\alpha-1)^2} \ ■$$

Practice Problems

Problem 31.1
Let X have the density function given by

$$f(x) = \begin{cases} 0.2 & -1 < x \leq 0 \\ 0.2 + cx & 0 < x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the value of c.
(b) Find $F(x)$.
(c) Find $\Pr(0 \leq x \leq 0.5)$.
(d) Find $E(X)$.

Problem 31.2
The density function of X is given by

$$f(x) = \begin{cases} a + bx^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Suppose that $E(X) = \frac{3}{5}$.
(a) Find a and b.
(b) Determine the cdf, $F(x)$, explicitly.

Problem 31.3
Compute $E(X)$ if X has the density function given by
(a)

$$f(x) = \begin{cases} \frac{1}{4}xe^{-\frac{x}{2}} & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(b)

$$f(x) = \begin{cases} c(1 - x^2) & -1 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

(c)

$$f(x) = \begin{cases} \frac{5}{x^2} & x > 5 \\ 0 & \text{otherwise.} \end{cases}$$

Problem 31.4
A continuous random variable has a pdf

$$f(x) = \begin{cases} 1 - \frac{x}{2} & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

Find the expected value and the variance.

Problem 31.5
Let X denote the lifetime (in years) of a computer chip. Let the probability density function be given by

$$f(x) = \begin{cases} 4(1+x)^{-5} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the mean and the standard deviation.
(b) What is the probability that a randomly chosen computer chip expires in less than a year?

Problem 31.6
Let X be a continuous random variable with pdf

$$f(x) = \begin{cases} \frac{1}{x} & 1 < x < e \\ 0 & \text{otherwise.} \end{cases}$$

Find $E(\ln X)$.

Problem 31.7
Let X have a cdf

$$F(x) = \begin{cases} 1 - \frac{1}{x^6} & x \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find $\text{Var}(X)$.

Problem 31.8
Let X have a pdf

$$f(x) = \begin{cases} 1 & 1 < x < 2 \\ 0 & \text{otherwise.} \end{cases}$$

Find the expected value and variance of $Y = X^2$.

Problem 31.9 ‡
Let X be a continuous random variable with density function

$$f(x) = \begin{cases} \frac{|x|}{10} & -2 \leq x \leq 4 \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the expected value of X.

Problem 31.10 ‡
An auto insurance company insures an automobile worth 15,000 for one year under a policy with a 1,000 deductible. During the policy year there is a 0.04 chance of partial damage to the car and a 0.02 chance of a total loss of the car. If there is partial damage to the car, the amount X of damage (in thousands) follows a distribution with density function

$$f(x) = \begin{cases} 0.5003e^{-0.5x} & 0 < x < 15 \\ 0 & \text{otherwise.} \end{cases}$$

What is the expected claim payment?

Problem 31.11 ‡
An insurance company's monthly claims are modeled by a continuous, positive random variable X, whose probability density function is proportional to $(1+x)^{-4}$, where $0 < x < \infty$ and 0 otherwise.
Determine the company's expected monthly claims.

Problem 31.12 ‡
A random variable X has the cumulative distribution function

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{x^2-2x+2}{2} & 1 \le x < 2 \\ 1 & x \ge 2. \end{cases}$$

Calculate the variance of X.

Problem 31.13 ‡
A company agrees to accept the highest of four sealed bids on a property. The four bids are regarded as four independent random variables with common cumulative distribution function

$$F(x) = \frac{1}{2}(1 + \sin \pi x), \quad \frac{3}{2} \le x \le \frac{5}{2}$$

and 0 otherwise. What is the expected value of the accepted bid?

Problem 31.14 ‡
An insurance policy on an electrical device pays a benefit of 4000 if the device fails during the first year. The amount of the benefit decreases by 1000 each successive year until it reaches 0 . If the device has not failed by the beginning of any given year, the probability of failure during that year is 0.4.
What is the expected benefit under this policy?

Problem 31.15
Let X have the density function

$$f(x) = \begin{cases} \lambda\frac{2x}{k^2} & 0 \le x \le k \\ 0 & \text{otherwise.} \end{cases}$$

For what value of k is the variance of X equal to 2?

Problem 31.16 ‡
A man purchases a life insurance policy on his 40^{th} birthday. The policy will pay 5000 only if he dies before his 50^{th} birthday and will pay 0 otherwise. The length of lifetime, in years, of a male born the same year as the insured has the cumulative distribution function

$$F(t) = \begin{cases} 1 - e^{\frac{1-1.1^t}{1000}}, & t > 0 \\ 0 & t \le 0 \end{cases}$$

Calculate the expected payment to the man under this policy.

32 Median, Mode, and Percentiles

In addition to the information provided by the mean and variance of a distribution, some other metrics such as the median, the mode, the percentile, and the quantile provide useful information.

Median of a Random Variable

In probability theory, **median** is described as the numerical value separating the higher half of a probability distribution, from the lower half. Thus, the median of a discrete random variable X is the number M such that $\Pr(X \leq M) \geq 0.50$ and $\Pr(X \geq M) \geq 0.50$.

Example 32.1

Given the pmf of a discrete random variable X.

x	0	1	2	3	4	5
$p(x)$	0.35	0.20	0.15	0.15	0.10	0.05

Find the median of X.

Solution.

Since $\Pr(X \leq 1) = 0.55$ and $\Pr(X \geq 1) = 0.65$, 1 is the median of X ■

In the case of a continuous random variable X, the median is the number M such that $\Pr(X \leq M) = \Pr(X \geq M) = 0.5$. Generally, M is found by solving the equation $F(M) = 0.5$ where F is the cdf of X.

Example 32.2

Let X be a continuous random variable with pdf $f(x) = \frac{1}{b-a}$ for $a < x < b$ and 0 otherwise. Find the median of X.

Solution.

We must find a number M such that $\int_a^M \frac{dx}{b-a} = 0.5$. This leads to the equation $\frac{M-a}{b-a} = 0.5$. Solving this equation we find $M = \frac{a+b}{2}$ ■

Remark 32.1

A discrete random variable might have many medians. For example, let X be the discrete random variable with pmf given by $p(x) = \left(\frac{1}{2}\right)^x,\ x = 1, 2, \cdots$ and 0 otherwise. Then any number $1 < M < 2$ satisfies $\Pr(X \leq M) = \Pr(X \geq M) = 0.5$.

Mode of a Random Variable

The **mode** is defined as the value that maximizes the probability mass function $p(x)$ (discrete case) or the probability density function $f(x)$ (continuous case.) In the discrete case, the mode is the value that is most likely to be sampled. In the continuous case, the mode is where $f(x)$ is at its peak.

Example 32.3
Let X be the discrete random variable with pmf given by $p(x) = \left(\frac{1}{2}\right)^x,\ x = 1, 2, \cdots$ and 0 otherwise. Find the mode of X.

Solution.
The value of x that maximizes $p(x)$ is $x = 1$. Thus, the mode of X is 1 ■

Example 32.4
Let X be the continuous random variable with pdf given by $f(x) = 0.75(1 - x^2)$ for $-1 \le x \le 1$ and 0 otherwise. Find the mode of X.

Solution.
The pdf is maximum for $x = 0$. Thus, the mode of X is 0 ■

Percentiles and Quantiles

In statistics, a percentile is the value of a variable below which a certain percent of observations fall. For example, if a score is in the 85^{th} percentile, it is higher than 85% of the other scores. For a random variable X and $0 < p < 1$, the 100p^{th} **percentile** (or the p^{th} **quantile**) is the number x such

$$\Pr(X < x) \le p \le \Pr(X \le x).$$

For a continuous random variable, this is the solution to the equation $F(x) = p$. The 25^{th} percentile is also known as the **first quartile**, the 50^{th} percentile as the median or second quartile, and the 75^{th} percentile as the third quartile.

Example 32.5
A loss random variable X has the density function

$$f(x) = \begin{cases} \frac{2.5(200)^{2.5}}{x^{3.5}} & x > 200 \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the difference between the 25^{th} and 75^{th} percentiles of X.

Solution.
First, the cdf is given by

$$F(x) = \int_{200}^{x} \frac{2.5(200)^{2.5}}{t^{3.5}} dt.$$

If Q_1 is the 25^{th} percentile then it satisfies the equation

$$F(Q_1) = \frac{1}{4}$$

or equivalently

$$1 - F(Q_1) = \frac{3}{4}.$$

This leads to

$$\frac{3}{4} = \int_{Q_1}^{\infty} \frac{2.5(200)^{2.5}}{t^{3.5}} dt = -\left(\frac{200}{t}\right)^{2.5}\Bigg|_{Q_1}^{\infty} = \left(\frac{200}{Q_1}\right)^{2.5}.$$

Solving for Q_1 we find $Q_1 = 200(4/3)^{0.4} \approx 224.4$. Similarly, the third quartile (i.e. 75^{th} percentile) is given by $Q_3 = 348.2$, The **interquartile range** (i.e., the difference between the 25^{th} and 75^{th} percentiles) is $Q_3 - Q_1 = 348.2 - 224.4 = 123.8$ ■

Example 32.6
Let X be the random variable with pdf $f(x) = \frac{1}{b-a}$ for $a < x < b$ and 0 otherwise. Find the p^{th} quantile of X.

Solution.
We have

$$p = \Pr(X \le x) = \int_a^x \frac{dt}{b-a} = \frac{x-a}{b-a}.$$

Solving this equation for x, we find $x = a + (b-a)p$ ■

Example 32.7
What percentile is 0.63 quantile?

Solution.
0.63 quantile is 63^{rd} percentile ■

Example 32.8 ‡
An insurance company sells an auto insurance policy that covers losses incurred by a policyholder, subject to a deductible of 100 . Losses incurred follow an exponential distribution with mean 300.
What is the 95th percentile of actual losses that exceed the deductible?

Solution.
The main difficulty here is the correct interpretation of the "95th percentile of actual losses that exceed the deductible." The proper interpretation involves a conditional probability: we seek the value x such that the conditional probability that the loss is at most x, given that it exceeds the deductible, is 0.95, i.e., $\Pr(X \leq x | X \geq 100) = 0.95$, where X denotes the loss, or equivalently

$$\frac{\Pr(100 \leq X \leq x)}{\Pr(X \geq 100)} = 0.95.$$

The above implies

$$\frac{F_X(x) - F_X(100)}{1 - F_X(100)} = 0.95$$

where $F_X(x) = 1 - e^{-\frac{1}{300}x}$. Simple algebra leads to the equation $0.95 = 1 - e^{\frac{1}{3}}e^{-\frac{x}{300}}$. Solving this equation for x we find $x = -300 \ln 0.05e^{-\frac{1}{3}} \approx 998.72$ ■

Practice Problems

Problem 32.1
Suppose the random variable X has pmf

$$\Pr(n) = \frac{1}{3}\left(\frac{2}{3}\right)^n, \quad n = 0, 1, 2, \cdots$$

Find the median and the 70^{th} percentile.

Problem 32.2
Suppose the random variable X has pdf

$$f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find the 50^{th} percentile.

Problem 32.3
Let Y be a continuous random variable with cumulative distribution function

$$F(y) = \begin{cases} 0 & y \leq a \\ 1 - e^{-\frac{1}{2}(y-a)^2} & \text{otherwise} \end{cases}$$

where a is a constant. Find the 75th percentile of Y.

Problem 32.4
Let X be a randon variable with density function

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find λ if the median of X is $\frac{1}{3}$.

Problem 32.5
People are dispersed on a linear beach with a density function $f(y) = 4y^3,\ 0 < y < 1$, and 0 elsewhere. An ice cream vendor wishes to locate her cart at the median of the locations (where half of the people will be on each side of her). Where will she locate her cart?

Problem 32.6 ‡
An automobile insurance company issues a one-year policy with a deductible of 500. The probability is 0.8 that the insured automobile has no accident and 0.0 that the automobile has more than one accident. If there is an accident, the loss before application of the deductible is exponentially distributed with mean 3000.
Calculate the 95^{th} percentile of the insurance company payout on this policy.

Problem 32.7
Using words, explain the meaning of $F(1120) = 0.2$ in terms of percentiles and quantiles.

Problem 32.8
Let X be a discrete random variable with pmf $p(n) = (n-1)(0.4)^2(0.6)^{n-2}$, $n \geq 2$ and 0 otherwise. Find the mode of X.

Problem 32.9
Let X be a continuous random variable with density function

$$f(x) = \begin{cases} \lambda\frac{1}{9}x(4-x) & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

Find the mode of X.

Problem 32.10
Find the pth quantile of the exponential distribution defined by the distribution function $F(x) = 1 - e^{-x}$ for $x \geq 0$ and 0 otherwise.

Problem 32.11
A continuous random variable has the pdf $f(x) = \frac{1}{2}e^{-|x|}$ for $x \in \mathbb{R}$. Find the pth quantile of X.

Problem 32.12
Let X be a loss random variable with cdf

$$F(x) = \begin{cases} 1 - \left(\frac{\theta}{\theta+x}\right)^{\alpha}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

The 10th percentile is $\theta - k$. The 90th percentile is $5\theta - 3k$. Determine the value of α.

Problem 32.13
Let X be a random variable with density function $f(x) = \frac{4x}{(1+x^2)^3}$ for $x > 0$ and 0 otherwise. Calculate the mode of X.

Problem 32.14
Let X be a random variable with pdf $f(x) = \left(\frac{3}{5000}\right)\left(\frac{5000}{x}\right)^4$ for $x > 5000$ and 0 otherwise. Determine the median of X.

Problem 32.15
Let X be a random variable with cdf

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^3}{27}, & 0 \leq x \leq 3 \\ 1, & x > 3. \end{cases}$$

Find the median of X.

Problem 32.16
A distribution has a pdf $f(x) = \frac{3}{x^4}$ for $x > 1$ and 0 otherwise. Calculate the 0.95^{th} quantile of this distribution.

33 The Uniform Distribution Function

The simplest continuous distribution is the uniform distribution. A continuous random variable X is said to be **uniformly** distributed over the interval $a \leq x \leq b$ if its pdf is given by

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

Since $F(x) = \int_{-\infty}^{x} f(t)dt$, the cdf is given by

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

Figure 33.1 presents a graph of $f(x)$ and $F(x)$.

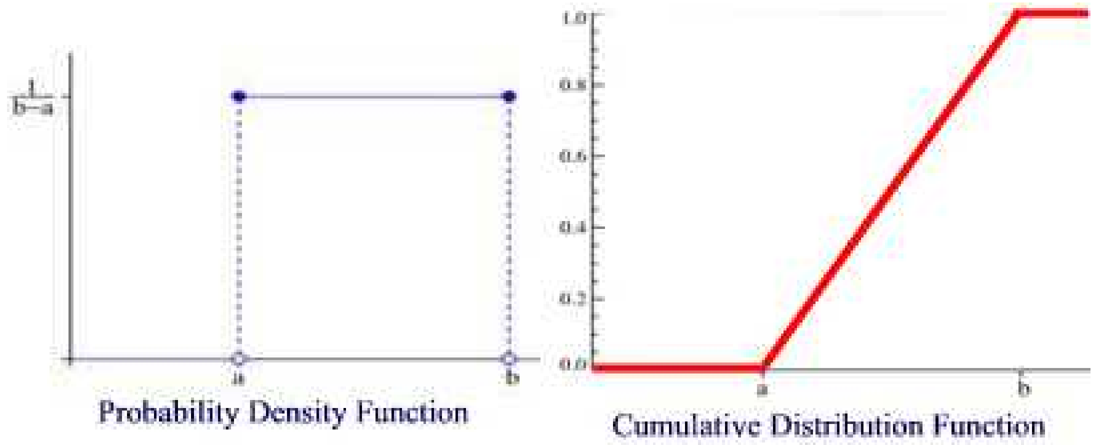


Figure 33.1

If $a = 0$ and $b = 1$ then X is called the **standard uniform** random variable.

Remark 33.1
The values at the two boundaries a and b are usually unimportant because they do not alter the value of the integral of $f(x)$ over any interval. Sometimes they are chosen to be zero, and sometimes chosen to be $\frac{1}{b-a}$. Our definition above assumes that $f(a) = f(b) = f(x) = \frac{1}{b-a}$. In the case $f(a) = f(b) = 0$ then the pdf becomes

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

Because the pdf of a uniform random variable is constant, if X is uniform, then the probability X lies in any interval contained in (a, b) depends only on the length of the interval-not location. That is, for any x and d such that $[x, x+d] \subseteq [a, b]$ we have

$$\int_x^{x+d} f(x)dx = \frac{d}{b-a}.$$

Hence uniformity is the continuous equivalent of a discrete sample space in which every outcome is equally likely.

Example 33.1
Find the survival function of a uniform distribution X on the interval $[a, b]$.

Solution.
The survival function is given by

$$S(x) = \begin{cases} 1 & \text{if } x \le a \\ \frac{b-x}{b-a} & \text{if } a < x < b \\ 0 & \text{if } x \ge b \ \blacksquare \end{cases}$$

Example 33.2
Let X be a continuous uniform random variable on $[0, 25]$. Find the pdf and cdf of X.

Solution.
The pdf is

$$f(x) = \begin{cases} \frac{1}{25} & 0 \le x \le 25 \\ 0 & \text{otherwise} \end{cases}$$

and the cdf is

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x}{25} & \text{if } 0 \le x \le 25 \\ 1 & \text{if } x > 25 \ \blacksquare \end{cases}$$

Example 33.3
Suppose that X has a uniform distribution on the interval $(0, a)$, where $a > 0$. Find $\Pr(X > X^2)$.

Solution.
If $a \leq 1$ then $\Pr(X > X^2) = \int_0^a \frac{1}{a}dx = 1$. If $a > 1$ then $\Pr(X > X^2) = \int_0^1 \frac{1}{a}dx = \frac{1}{a}$. Thus, $\Pr(X > X^2) = \min\{1, \frac{1}{a}\}$ ■

The expected value of X is

$$\begin{aligned}E(X) =& \int_a^b xf(x) = \int_a^b \frac{x}{b-a}dx\\ =& \left.\frac{x^2}{2(b-a)}\right|_a^b = \frac{b^2-a^2}{2(b-a)}\\ =&\frac{a+b}{2}\end{aligned}$$

and so the expected value of a uniform random variable is halfway between a and b.
The second moment about the origin is

$$\begin{aligned}E(X^2) =& \int_a^b \frac{x^2}{b-a}dx = \left.\frac{x^3}{3(b-a)}\right|_a^b\\ =&\frac{b^3-a^3}{3(b-a)} = \frac{a^2+b^2+ab}{3}.\end{aligned}$$

The variance of X is

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{a^2+b^2+ab}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}.$$

Practice Problems

Problem 33.1
Let X be the total time to process a passport application by the state department. It is known that X is uniformly distributed between 3 and 7 weeks.
(a) Find $f(x)$.
(b) What is the probability that an application will be processed in fewer than 3 weeks ?
(c) What is the probability that an application will be processed in 5 weeks or less ?

Problem 33.2
In a sushi bar, customers are charged for the amount of sushi they consume. Suppose that the amount of sushi consumed is uniformly distributed between 5 ounces and 15 ounces. Let X be the random variable representing a plate filling weight.
(a) Find the probability density function of X.
(b) What is the probability that a customer will take between 12 and 15 ounces of sushi?
(c) Find $E(X)$ and Var(X).

Problem 33.3
Suppose that X has a uniform distribution over the interval (0, 1). Find
(a) $F(x)$.
(b) Show that $\Pr(a \leq X \leq a+b)$ for $a, b \geq 0,\ a+b \leq 1$ depends only on b.

Problem 33.4
Let X be uniform on (0,1). Compute $E(X^n)$ where n is a positive integer.

Problem 33.5
Let X be a uniform random variable on the interval (1,2) and let $Y = \frac{1}{X}$. Find $E[Y]$.

Problem 33.6
A commuter train arrives at a station at some time that is uniformly distributes between 10:00 AM and 10:30 AM. Let X be the waiting time (in minutes) for the train. What is the probability that you will have to wait longer than 10 minutes?

Problem 33.7 ‡
An insurance policy is written to cover a loss, X, where X has a uniform distribution on [0, 1000].
At what level must a deductible be set in order for the expected payment to be 25% of what it would be with no deductible?

Problem 33.8 ‡
The warranty on a machine specifies that it will be replaced at failure or age 4, whichever occurs first. The machine's age at failure, X, has density function

$$f(x) = \begin{cases} \frac{1}{5} & 0 < x < 5 \\ 0 & \text{otherwise} \end{cases}$$

Let Y be the age of the machine at the time of replacement. Determine the variance of Y.

Problem 33.9 ‡
The owner of an automobile insures it against damage by purchasing an insurance policy with a deductible of 250 . In the event that the automobile is damaged, repair costs can be modeled by a uniform random variable on the interval $(0, 1500)$.
Determine the standard deviation of the insurance payment in the event that the automobile is damaged.

Problem 33.10
Let X be a random variable distributed uniformly over the interval $[-1, 1]$.
(a) Compute $E(e^{-X})$.
(b) Compute $\mathrm{Var}(e^{-X})$.

Problem 33.11
Let X be a random variable with a continuous uniform distribution on the interval $(1, a), a > 1$. If $E(X) = 6\mathrm{Var}(X)$, what is the value of a?

Problem 33.12
Let X be a random variable with a continuous uniform distribution on the interval $(0, 10)$. What is $\Pr(X + \frac{10}{X} > 7)$?

34 Normal Random Variables

A **normal** random variable with parameters μ and σ^2 has a pdf

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

This density function is a bell-shaped curve that is symmetric about μ (See Figure 34.1).

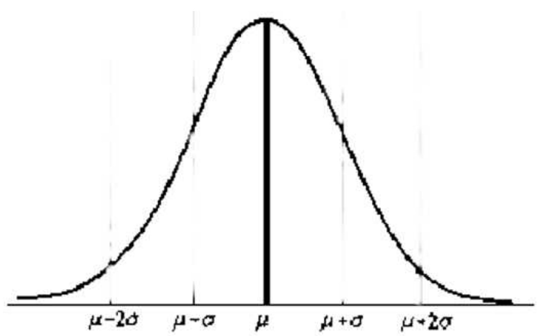


Figure 34.1

The normal distribution is used to model phenomenon such as a person's height at a certain age or the measurement error in an experiment. Observe that the distribution is symmetric about the point $\mu-$hence the experiment outcome being modeled should be equaly likely to assume points above μ as points below μ. The normal distribution is probably the most important distribution because of a result we will disuss in Section 51, known as the *central limit theorem.*
To prove that the given $f(x)$ is indeed a pdf we must show that the area under the normal curve is 1. That is,

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

First note that using the substitution $y = \frac{x-\mu}{\sigma}$ we have

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

Toward this end, let $I = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy$. Then

$$\begin{aligned} I^2 =& \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dxdy \\ =& \int_0^{\infty}\int_0^{2\pi} e^{-\frac{r^2}{2}} rd\theta dr = 2\pi \int_0^{\infty} re^{-\frac{r^2}{2}} dr = 2\pi. \end{aligned}$$

Thus, $I = \sqrt{2\pi}$ and the result is proved. Note that in the process above, we used the polar substitution $x = r\cos\theta, y = r\sin\theta$, and $dydx = rdrd\theta$.

Example 34.1
Let X be a normal random variable with mean 950 and standard deviation 10. Find $\Pr(947 \leq X \leq 950)$.

Solution.
We have

$$\Pr(947 \leq X \leq 950) = \frac{1}{10\sqrt{2\pi}} \int_{947}^{950} e^{-\frac{(x-950)^2}{200}} dx \approx 0.118$$

where the value of the integral is found by using a calculator ■

Theorem 34.1
If X is a normal distribution with parameters (μ, σ^2) then $Y = aX + b$ is a normal distribution with paramaters $(a\mu + b, a^2\sigma^2)$.

Proof.
We prove the result when $a > 0$. The proof is similar for $a < 0$. Let F_Y denote the cdf of Y. Then

$$\begin{aligned} F_Y(x) =& \Pr(Y \leq x) = \Pr(aX + b \leq x) \\ =& P\left(X \leq \frac{x-b}{a}\right) = F_X\left(\frac{x-b}{a}\right). \end{aligned}$$

Differentiating both sides to obtain

$$\begin{aligned} f_Y(x) =& \frac{1}{a} f_X\left(\frac{x-b}{a}\right) = \frac{1}{\sqrt{2\pi}a\sigma} exp\left[-(\frac{x-b}{a} - \mu)^2/(2\sigma^2)\right] \\ =& \frac{1}{\sqrt{2\pi}a\sigma} exp\left[-(x-(a\mu+b))^2/2(a\sigma)^2\right] \end{aligned}$$

which shows that Y is normal with parameters $(a\mu + b, a^2\sigma^2)$ ■
Note that if $Z = \frac{X-\mu}{\sigma}$ then this is a normal distribution with parameters (0,1). Such a random variable is called the **standard normal** random variable.

Theorem 34.2
If X is a normal random variable with parameters (μ, σ^2) then
(a) $E(X) = \mu$
(b) $\text{Var}(X) = \sigma^2$.

Proof.
(a) Let $Z = \frac{X-\mu}{\sigma}$ be the standard normal distribution. Then

$E(Z) = \int_{-\infty}^{\infty} x f_Z(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} dx = -\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} = 0.$

Thus,

$$E(X) = E(\sigma Z + \mu) = \sigma E(Z) + \mu = \mu.$$

(b)

$$\text{Var}(Z) = E(Z^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx.$$

Using integration by parts with $u = x$ and $dv = xe^{-\frac{x^2}{2}}$ we find

$\text{Var(Z)} = \frac{1}{\sqrt{2\pi}} \left[-xe^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1.$

Thus,

$$\text{Var}(X) = \text{Var}(\sigma Z + \mu) = \sigma^2 \text{Var}(Z) = \sigma^2 \blacksquare$$

Figure 34.2 shows different normal curves with the same μ and different σ.

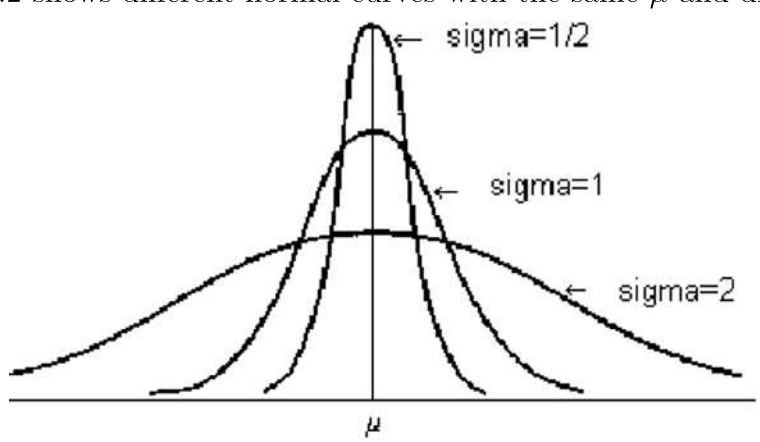


Figure 34.2

It is traditional to denote the cdf of Z by $\Phi(x)$. That is,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy.$$

Now, since $f_Z(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $f_Z(x)$ is an even function. This implies that $\Phi'(-x) = \Phi'(x)$. Integrating we find that $\Phi(x) = -\Phi(-x) + C$. Letting $x = 0$ we find that $C = 2\Phi(0) = 2(0.5) = 1$. Thus,

$$\Phi(x) = 1 - \Phi(-x), \quad -\infty < x < \infty. \tag{34.1}$$

This implies that

$$\Pr(Z \le -x) = \Pr(Z > x).$$

Now, $\Phi(x)$ is the area under the standard curve to the left of x. The values of $\Phi(x)$ for $x \ge 0$ are given in Table 34.1 below. Equation 34.1 is used for $x < 0$.

Example 34.2
Let X be a normal random variable with parameters $\mu = 24$ and $\sigma_X^2 = 9$.
(a) Find $\Pr(X > 27)$ using Table 34.1.
(b) Solve $S(x) = 0.05$ where $S(x)$ is the survival function of X.

Solution.
(a) The desired probability is given by

$$\begin{aligned}\Pr(X > 27) =& P\left(\frac{X-24}{3} > \frac{27-24}{3}\right) = \Pr(Z > 1)\\ =& 1 - \Pr(Z \le 1) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587.\end{aligned}$$

(b) The equation $\Pr(X > x) = 0.05$ is equivalent to $\Pr(X \le x) = 0.95$. Note that

$$\Pr(X \le x) = P\left(\frac{X-24}{3} < \frac{x-24}{3}\right) = P\left(Z < \frac{x-24}{3}\right) = 0.95$$

From Table 34.1 we find $\Pr(Z \le 1.65) = 0.95$. Thus, we set $\frac{x-24}{3} = 1.65$ and solve for x we find $x = 28.95$ ■

From the above example, we see that probabilities involving normal random variables can be reduced to the ones involving standard normal variable. For example

$$\Pr(X \le a) = P\left(\frac{X-\mu}{\sigma} \le \frac{a-\mu}{\sigma}\right) = \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Example 34.3
Let X be a normal random variable with parameters μ and σ^2. Find
(a)$\Pr(\mu - \sigma \le X \le \mu + \sigma)$.
(b)$\Pr(\mu - 2\sigma \le X \le \mu + 2\sigma)$.
(c)$\Pr(\mu - 3\sigma \le X \le \mu + 3\sigma)$.

Solution.
(a) We have

$$\begin{aligned}\Pr(\mu - \sigma \le X \le \mu + \sigma) =& \Pr(-1 \le Z \le 1)\\ =& \Phi(1) - \Phi(-1)\\ =& 2(0.8413) - 1 = 0.6826.\end{aligned}$$

Thus, 68.26% of all possible observations lie within one standard deviation to either side of the mean.
(b) We have

$$\begin{aligned}\Pr(\mu - 2\sigma \le X \le \mu + 2\sigma) =& \Pr(-2 \le Z \le 2) = \Phi(2) - \Phi(-2)\\ =& 2(0.9772) - 1 = 0.9544.\end{aligned}$$

Thus, 95.44% of all possible observations lie within two standard deviations to either side of the mean.
(c) We have

$$\begin{aligned}\Pr(\mu - 3\sigma \le X \le \mu + 3\sigma) =& \Pr(-3 \le Z \le 3) = \Phi(3) - \Phi(-3)\\ =& 2(0.9987) - 1 = 0.9974.\end{aligned}$$

Thus, 99.74% of all possible observations lie within three standard deviations to either side of the mean. See Figure 34.3 ■

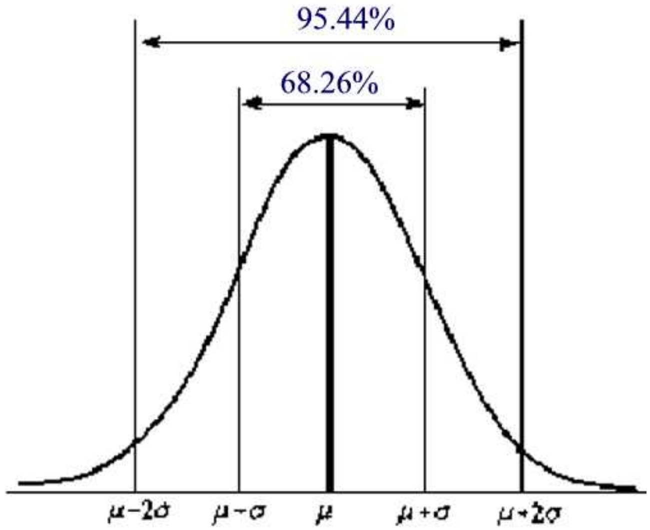


Figure 34.3

Table 34.1: Area under the Standard Normal Curve from $-\infty$ to x

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Practice Problems

Problem 34.1
The scores on a statistics test are Normally distributed with parameters $\mu = 80$ and $\sigma^2 = 196$. Find the probability that a randomly chosen score is
(a) no greater than 70
(b) at least 95
(c) between 70 and 95.
(d) Approximately, what is the raw score corresponding to a percentile score of 72%?

Problem 34.2
Let X be a normal random variable with parameters $\mu = 0.381$ and $\sigma^2 = 0.031^2$. Compute the following:
(a) $\Pr(X > 0.36)$.
(b) $\Pr(0.331 < X < 0.431)$.
(c) $\Pr(|X - .381| > 0.07)$.

Problem 34.3
Assume the time required for a cyclist to travel a distance d follows a normal distribution with mean 4 minutes and variance 4 seconds.
(a) What is the probability that this cyclist with travel the distance in less than 4 minutes?
(b) What is the probability that this cyclist will travel the distance in between 3min55sec and 4min5sec?

Problem 34.4
It has been determined that the lifetime of a certain light bulb has a normal distribution with $\mu = 2000$ hours and $\sigma = 200$ hours.
(a) Find the probability that a bulb will last between 2000 and 2400 hours.
(b) What is the probability that a light bulb will last less than 1470 hours?

Problem 34.5
Let X be a normal random variable with mean 100 and standard deviation 15. Find $\Pr(X > 130)$ given that $\Phi(2) = .9772$.

Problem 34.6
The lifetime X of a randomly chosen battery is normally distributed with mean 50 and standard devaition 5.
(a) Find the probability that the battery lasts at least 42 hours.
(b) Find the probability that the battery will last between 45 to 60 hours.

Problem 34.7 ‡
For Company A there is a 60% chance that no claim is made during the coming year. If one or more claims are made, the total claim amount is normally distributed with mean 10,000 and standard deviation 2,000 .
For Company B there is a 70% chance that no claim is made during the coming year. If one or more claims are made, the total claim amount is normally distributed with mean 9,000 and standard deviation 2,000 .
Assuming that the total claim amounts of the two companies are independent, what is the probability that, in the coming year, Company B's total claim amount will exceed Company A's total claim amount?

Problem 34.8
Let X be a normal random variable with $\Pr(X < 500) = 0.5$ and $\Pr(X > 650) = 0.0227$. Find the standard deviation of X.

Problem 34.9
Suppose that X is a normal random variable with parameters $\mu = 5, \sigma^2 = 49$. Using the table of the normal distribution , compute: (a) $\Pr(X > 5.5)$, (b) $\Pr(4 < X < 6.5)$, (c) $\Pr(X < 8)$, (d) $\Pr(|X - 7| \geq 4)$.

Problem 34.10
Let X be a normal random variable with mean 1 and variance 4. Find $\Pr(X^2 - 2X \leq 8)$.

Problem 34.11
Let X be a normal random variable with mean 360 and variance 16.
(a) Calculate $\Pr(X < 355)$.
(b) Suppose the variance is kept at 16 but the mean is to be adjusted so that $\Pr(X < 355) = 0.025$. Find the adjusted mean.

Problem 34.12
The length of time X (in minutes) it takes to go from your home to donwtown is normally distributed with $\mu = 30$ minutes and $\sigma_X = 5$ minutes. What is the latest time that you should leave home if you want to be over 99% sure of arriving in time for a job interview taking place in downtown at 2pm?

35 The Normal Approximation to the Binomial Distribution

When the number of trials in a binomial distribution is very large, the use of the probability distribution formula $p(x) =_n C_x p^x q^{n-x}$ becomes tedious. An attempt was made to approximate this distribution for large values of n. The approximating distribution is the normal distribution.
Historically, the normal distribution was discovered by De Moivre as an approximation to the binomial distribution. The result is the so-called *De Moivre-Laplace theorem.*

Theorem 35.1
Let S_n denote the number of successes that occur with n independent Bernoulli trials, each with probability p of success. Then, for $a < b$,

$$\lim_{n\to\infty} P\left[a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right] = \Phi(b) - \Phi(a)$$

where $\Phi(x)$ is the cdf of the standard normal distribution.

Proof.
This result is a special case of the central limit theorem, which will be discussed in Section 51. Consequently, we will defer the proof of this result until then ■

Remark 35.1
How large should n be so that a normal approximation to the binomial distribution is adequate? A rule-of-thumb for the normal distribution to be a good approximation to the binomial distribution is to have $np > 5$ and $nq > 5$.

Remark 35.2 *(continuity correction)*
Suppose we are approximating a binomial random variable with a normal random variable. Say we want to find $\Pr(8 \leq X \leq 10)$ where X is a binomial distribution. According to Figure 35.1, the probability in question is the area of the two rectangles centered at 8 and 9. When using the normal distribution to approximate the binomial distribution, the area under the pdf from 7.5 to 10.5 must be found. That is,

$$\Pr(8 \leq X \leq 10) = \Pr(7.5 \leq N \leq 10.5)$$

where N is the corresponding normal variable. In practice, then, we apply a *continuity correction,* when approximating a discrete random variable with a continuous random variable.

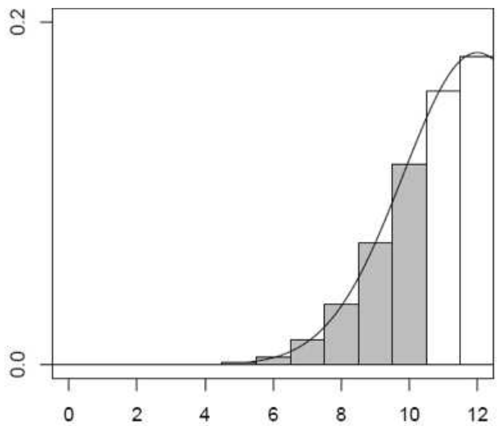


Figure 35.1

Example 35.1
In a box of 100 light bulbs, 10 are found to be defective. What is the probability that the number of defectives exceeds 13?

Solution.
Let X be the number of defective items. Then X is binomial with $n = 100$ and $p = 0.1$. Since $np = 10 > 5$ and $nq = 9 > 5$ we can use the normal approximation to the binomial with $\mu = np = 10$ and $\sigma^2 = np(1-p) = 9$. We want $\Pr(X > 13)$. Using continuity correction we find

$$\begin{aligned}\Pr(X > 13) =&\Pr(X \geq 14)\\ =&\Pr(\frac{X-10}{\sqrt{9}} \geq \frac{13.5-10}{\sqrt{9}})\\ \approx&1 - \Phi(1.17) = 1 - 0.8790 = 0.121 \blacksquare\end{aligned}$$

Example 35.2
In a small town, it was found that out of every 6 people 1 is left-handed. Consider a random sample of 612 persons from the town, estimate the probability that the number of lefthanded persons is strictly between 90 and 150.

Solution.
Let X be the number of left-handed people in the sample. Then X is a binomial random variable with $n = 612$ and $p = \frac{1}{6}$. Since $np = 102 > 5$ and $n(1-p) = 510 > 5$ we can use the normal approximation to the binomial with $\mu = np = 102$ and $\sigma^2 = np(1-p) = 85$. Using continuity correction we find

$$\begin{aligned}\Pr(90 < X < 150) =& \Pr(91 \le X \le 149) = \\ =& \Pr\left(\frac{90.5 - 102}{\sqrt{85}} \le \frac{X - 102}{\sqrt{85}} \le \frac{149.5 - 102}{\sqrt{85}}\right) \\ =& \Pr(-1.25 \le Z \le 5.15) \approx 0.8943 \blacksquare\end{aligned}$$

Example 35.3
There are 90 students in a statistics class. Suppose each student has a standard deck of 52 cards of his/her own, and each of them selects 13 cards at random without replacement from his/her own deck independent of the others. What is the chance that there are more than 50 students who got at least 2 aces ?

Solution.
Let X be the number of students who got at least 2 aces or more, then clearly X is a binomial random variable with $n = 90$ and

$$p = \frac{{}_4C_2 \cdot {}_{48}C_{11}}{{}_{52}C_{13}} + \frac{{}_4C_3 \cdot {}_{48}C_{10}}{{}_{52}C_{13}} + \frac{{}_4C_4 \cdot {}_{48}C_9}{{}_{52}C_{13}} \approx 0.2573$$

Since $np \approx 23.157 > 5$ and $n(1-p) \approx 66.843 > 5$, X can be approximated by a normal random variable with $\mu = 23.157$ and $\sigma = \sqrt{np(1-p)} \approx 4.1473$. Thus,

$$\begin{aligned}\Pr(X > 50) =& 1 - \Pr(X \le 50) = 1 - \Phi\left(\frac{50.5 - 23.157}{4.1473}\right) \\ \approx& 1 - \Phi(6.59) \blacksquare\end{aligned}$$

Practice Problems

Problem 35.1
Suppose that 25% of all the students who took a given test fail. Let X be the number of students who failed the test in a random sample of 50.
(a) What is the probability that the number of students who failed the test is at most 10?
(b) What is the probability that the number of students who failed the test is between 5 and 15 inclusive?

Problem 35.2
A vote on whether to allow the use of medical marijuana is being held. A polling company will survey 200 individuals to measure support for the new law. If in fact 53% of the population oppose the new law, use the normal approximation to the binomial, with a continuity correction, to approximate the probability that the poll will show a majority in favor?

Problem 35.3
A company manufactures 50,000 light bulbs a day. For every 1,000 bulbs produced there are 50 bulbs defective. Consider testing a random sample of 400 bulbs from today's production. Find the probability that the sample contains
(a) At least 14 and no more than 25 defective bulbs.
(b) At least 33 defective bulbs.

Problem 35.4
Suppose that the probability of a family with two children is 0.25 that the children are boys. Consider a random sample of 1,000 families with two children. Find the probability that at most 220 families have two boys.

Problem 35.5
A survey shows that 10% of the students in a college are left-handed. In a random sample of 818, what is the probability that at most 100 students are left-handed?

36 Exponential Random Variables

An **exponential** random variable with parameter $\lambda > 0$ is a random variable with pdf

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

Note that

$$\int_0^\infty \lambda e^{-\lambda x} dx = -e^{-\lambda x}\Big|_0^\infty = 1.$$

The graph of the probability density function is shown in Figure 36.1

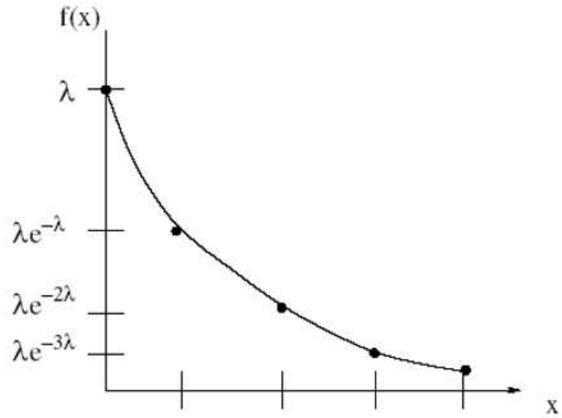


Figure 36.1

Exponential random variables are often used to model arrival times, waiting times, and equipment failure times.
The expected value of X can be found using integration by parts with $u = x$ and $dv = \lambda e^{-\lambda x} dx$:

$$\begin{aligned} E(X) &= \int_0^\infty x\lambda e^{-\lambda x} dx \\ &= \left[-xe^{-\lambda x}\right]_0^\infty + \int_0^\infty e^{-\lambda x} dx \\ &= \left[-xe^{-\lambda x}\right]_0^\infty + \left[-\frac{1}{\lambda}e^{-\lambda x}\right]_0^\infty \\ &= \frac{1}{\lambda} \end{aligned}$$

Furthermore, using integration by parts again, we may also obtain that

$$\begin{aligned}E(X^2) &= \int_0^\infty \lambda x^2 e^{-\lambda x}dx = \int_0^\infty x^2 d(-e^{-\lambda x})\\&= \left[-x^2e^{-\lambda x}\right]_0^\infty + 2\int_o^\infty xe^{-\lambda x}dx\\&= \frac{2}{\lambda^2}\end{aligned}$$

Thus,

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Example 36.1
The time between calls received by a 911 operator has an exponential distribution with an average of 3 calls per hour.
(a) Find the expected time between calls.
(b) Find the probability that the next call is received within 5 minutes.

Solution.
Let X denote the time (in hours) between calls. We are told that $\lambda = 3$.
(a) We have $E(X) = \frac{1}{\lambda} = \frac{1}{3}$.
(b) $\Pr(X < \frac{1}{12}) = \int_0^{\frac{1}{12}} 3e^{-3x}dx \approx 0.2212$ ■

Example 36.2
The time between hits to my website is an exponential distribution with an average of 2 minutes between hits. Suppose that a hit has just occurred to my website. Find the probability that the next hit won't happen within the next 5 minutes.

Solution.
Let X denote the time (in minutes) between two hits. Then X is an exponential distribution with paramter $\lambda = \frac{1}{2} = 0.5$. Thus,

$$\Pr(X > 5) = \int_5^\infty 0.5e^{-0.5x}dx \approx 0.082085 \blacksquare$$

The cumulative distribution function of an exponential random variable X is given by

$$F(x) = \Pr(X \le x) = \int_0^x \lambda e^{-\lambda u}du = -e^{-\lambda u}\,|_0^x = 1 - e^{-\lambda x}$$

for $x \ge 0$, and 0 otherwise.

Example 36.3
Suppose that the waiting time (in minutes) at a post office is an exponential random variable with mean 10 minutes. If someone arrives immediately ahead of you at the post office, find the probability that you have to wait
(a) more than 10 minutes
(b) between 10 and 20 minutes.

Solution.
Let X be the time you must wait in line at the post office. Then X is an exponential random variable with parameter $\lambda = 0.1$.
(a) We have $\Pr(X > 10) = 1 - F(10) = 1 - (1 - e^{-1}) = e^{-1} \approx 0.3679$.
(b) We have $\Pr(10 \leq X \leq 20) = F(20) - F(10) = e^{-1} - e^{-2} \approx 0.2325$ ■

The most important property of the exponential distribution is known as the **memoryless** property:

$$\Pr(X > s + t | X > s) = \Pr(X > t), \quad s, t \geq 0.$$

This says that the probability that we have to wait for an additional time t (and therefore a total time of $s + t$) given that we have already waited for time s is the same as the probability at the start that we would have had to wait for time t. So the exponential distribution "forgets" that it is larger than s.
To see why the memoryless property holds, note that for all $t \geq 0$, we have

$$\Pr(X > t) = \int_t^\infty \lambda e^{-\lambda x} dx = -e^{-\lambda x} |_t^\infty = e^{-\lambda t}.$$

It follows that

$$\begin{aligned}
\Pr(X > s + t | X > s) =& \frac{\Pr(X > s + t \text{ and } X > s)}{\Pr(X > s)} \\
=& \frac{\Pr(X > s + t)}{\Pr(X > s)} \\
=& \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\
=& e^{-\lambda t} = \Pr(X > t)
\end{aligned}$$

Example 36.4
Suppose that the time X (in hours) required to repair a car has an exponential

distribution with parameter $\lambda = 0.25$. Find
(a) the cumulative distribution function of X.
(b) $\Pr(X > 4)$.
(c) $\Pr(X > 10|X > 8)$.

Solution.
(a) It is easy to see that the cumulative distribution function is

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{4}} & x \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

(b) $\Pr(X > 4) = 1 - \Pr(X \leq 4) = 1 - F(4) = 1 - (1 - e^{-\frac{4}{4}}) = e^{-1} \approx 0.368$.
(c) By the memoryless property, we find

$$\begin{aligned} \Pr(X > 10|X > 8) =& \Pr(X > 8 + 2|X > 8) = \Pr(X > 2) \\ =& 1 - \Pr(X \leq 2) = 1 - F(2) \\ =& 1 - (1 - e^{-\frac{1}{2}}) = e^{-\frac{1}{2}} \approx 0.6065 \blacksquare \end{aligned}$$

Example 36.5
The time between hits to my website is an exponential distribution with an average of 5 minutes between hits.
(a) What is the probability that there are no hits in a 20-minute period?
(b) What is the probability that the first observed hit occurs between 15 and 20 minutes?
(c) Given that there are no hits in the first 5 minutes observed, what is the probability that there are no hits in the next 15 minutes?

Solution.
Let X denote the time between two hits. Then, X is an exponential random variable with $\mu = \frac{1}{E(X}) = \frac{1}{5} = 0.2$ hit/minute.
(a)

$$\Pr(X > 20) = \int_{20}^{\infty} 0.2e^{-0.2x}dx = -e^{-0.2x}\Big|_{20}^{\infty} = e^{-4} \approx 0.01831.$$

(b)

$$\Pr(15 < X < 20) = \int_{15}^{20} 0.2e^{-0.2x}dx = -e^{-0.2x}\Big|_{15}^{20} \approx 0.03147.$$

(c) By the memoryless property, we have

$$\Pr(X > 15{+}5|X > 5) = \Pr(X > 15) = \int_{15}^{\infty} 0.2e^{-0.2x}dx = -e^{-0.2x}\Big|_{15}^{\infty} \approx 0.04979 \blacksquare$$

The exponential distribution is the only named continuous distribution that possesses the memoryless property. To see this, suppose that X is memoryless continuous random variable. Let $g(x) = \Pr(X > x)$. Since X is memoryless, we have

$$\Pr(X > t) = \Pr(X > s+t|X > s) = \frac{\Pr(X > s+t \text{ and } X > s)}{\Pr(X > s)} = \frac{\Pr(X > s+t)}{\Pr(X > s)}$$

and this implies

$$\Pr(X > s+t) = \Pr(X > s)\Pr(X > t)$$

Hence, g satisfies the equation

$$g(s+t) = g(s)g(t).$$

Theorem 36.1
The only solution to the functional equation $g(s+t) = g(s)g(t)$ which is continuous from the right is $g(x) = e^{-\lambda x}$ for some $\lambda > 0$.

Proof.
Let $c = g(1)$. Then $g(2) = g(1+1) = g(1)^2 = c^2$ and $g(3) = c^3$ so by simple induction we can show that $g(n) = c^n$ for any positive integer n.
Now, let n be a positive integer, then $\left[g\left(\frac{1}{n}\right)\right]^n = g\left(\frac{1}{n}\right)g\left(\frac{1}{n}\right)\cdots g\left(\frac{1}{n}\right) = g\left(\frac{n}{n}\right) = c$. Thus, $g\left(\frac{1}{n}\right) = c^{\frac{1}{n}}$.
Next, let m and n be two positive integers. Then $g\left(\frac{m}{n}\right) = g\left(m\cdot\frac{1}{n}\right) = g\left(\frac{1}{n}+\frac{1}{n}+\cdots+\frac{1}{n}\right) = \left[g\left(\frac{1}{n}\right)\right]^m = c^{\frac{m}{n}}$.
Now, if t is a positive real number then we can find a sequence t_n of positive rational numbers such that $\lim_{n\to\infty} t_n = t$. (This is known as the density property of the real numbers and is a topic discussed in a real analysis course). Since $g(t_n) = c^{t_n}$, the right-continuity of g implies $g(t) = c^t,\ t \geq 0$.
Finally, let $\lambda = -\ln c$. Since $0 < c < 1$, we have $\lambda > 0$. Moreover, $c = e^{-\lambda}$ and therefore $g(t) = e^{-\lambda t},\ t \geq 0$ ■

It follows from the previous theorem that $F(x) = \Pr(X \leq x) = 1 - e^{-\lambda x}$ and hence $f(x) = F'(x) = \lambda e^{-\lambda x}$ which shows that X is exponentially distributed.

Example 36.6
Very often, credit card customers are placed on hold when they call for

inquiries. Suppose the amount of time until a service agent assists a customer has an exponential distribution with mean 5 minutes. Given that a customer has already been on hold for 2 minutes, what is the probability that he/she will remain on hold for a total of more than 5 minutes?

Solution.
Let X represent the total time on hold. Then X is an exponential random variable with $\lambda = \frac{1}{5}$. Thus,

$$\Pr(X > 3 + 2|X > 2) = \Pr(X > 3) = 1 - F(3) = e^{-\frac{3}{5}} \blacksquare$$

Practice Problems

Problem 36.1
Let X have an exponential distribution with a mean of 40. Compute $\Pr(X < 36)$.

Problem 36.2
Let X be an exponential function with mean equals to 5. Graph $f(x)$ and $F(x)$.

Problem 36.3
A continuous random variable X has the following pdf:

$$f(x) = \begin{cases} \frac{1}{100}e^{-\frac{x}{100}} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Compute $\Pr(0 \leq X \leq 50)$.

Problem 36.4
Let X be an exponential random variable with mean equals to 4. Find $\Pr(X \leq 0.5)$.

Problem 36.5
The life length X (in years) of a dvd player is exponentially distributed with mean 5 years. What is the probability that a more than 5-year old dvd would still work for more than 3 years?

Problem 36.6
Suppose that the spending time X (in minutes) of a customer at a bank has an exponential distribution with mean 3 minutes.
(a) What is the probability that a customer spends more than 5 minutes in the bank?
(b) Under the same conditions, what is the probability of spending between 2 and 4 minutes?

Problem 36.7
The waiting time X (in minutes) of a train arrival to a station has an exponential distribution with mean 3 minutes.
(a) What is the probability of having to wait 6 or more minutes for a train?
(b) What is the probability of waiting between 4 and 7 minutes for a train?
(c) What is the probability of having to wait at least 9 more minutes for the train given that you have already waited 3 minutes?

Problem 36.8 ‡
Ten years ago at a certain insurance company, the size of claims under homeowner insurance policies had an exponential distribution. Furthermore, 25% of claims were less than \$1000. Today, the size of claims still has an exponential distribution but, owing to inflation, every claim made today is twice the size of a similar claim made 10 years ago. Determine the probability that a claim made today is less than \$1000.

Problem 36.9
The lifetime (in hours) of a battery installed in a radio is an exponentially distributed random variable with parameter $\lambda = 0.01$. What is the probability that the battery is still in use one week after it is installed?

Problem 36.10 ‡
The number of days that elapse between the beginning of a calendar year and the moment a high-risk driver is involved in an accident is exponentially distributed. An insurance company expects that 30% of high-risk drivers will be involved in an accident during the first 50 days of a calendar year.
What portion of high-risk drivers are expected to be involved in an accident during the first 80 days of a calendar year?

Problem 36.11 ‡
The lifetime of a printer costing 200 is exponentially distributed with mean 2 years. The manufacturer agrees to pay a full refund to a buyer if the printer fails during the first year following its purchase, and a one-half refund if it fails during the second year.
If the manufacturer sells 100 printers, how much should it expect to pay in refunds?

Problem 36.12 ‡
A device that continuously measures and records seismic activity is placed in a remote region. The time, T, to failure of this device is exponentially distributed with mean 3 years. Since the device will not be monitored during its first two years of service, the time to discovery of its failure is $X = \max(T, 2)$.
Determine $E[X]$.

Problem 36.13 ‡
A piece of equipment is being insured against early failure. The time from

purchase until failure of the equipment is exponentially distributed with mean 10 years. The insurance will pay an amount x if the equipment fails during the first year, and it will pay $0.5x$ if failure occurs during the second or third year. If failure occurs after the first three years, no payment will be made. At what level must x be set if the expected payment made under this insurance is to be 1000 ?

Problem 36.14 ‡
An insurance policy reimburses dental expense, X, up to a maximum benefit of 250 . The probability density function for X is:

$$f(x) = \left\{ \begin{array}{cc} ce^{-0.004x} & x \geq 0 \\ 0 & \text{otherwise} \end{array} \right.$$

where c is a constant. Calculate the median benefit for this policy.

Problem 36.15 ‡
The time to failure of a component in an electronic device has an exponential distribution with a median of four hours.
Calculate the probability that the component will work without failing for at least five hours.

Problem 36.16
Let X be an exponential random variable such that $\Pr(X \leq 2) = 2\Pr(X > 4)$. Find the variance of X.

Problem 36.17 ‡
The cumulative distribution function for health care costs experienced by a policyholder is modeled by the function

$$F(x) = \left\{ \begin{array}{ll} 1 - e^{-\frac{x}{100}}, & \text{for } x > 0 \\ 0, & \text{otherwise.} \end{array} \right.$$

The policy has a deductible of 20. An insurer reimburses the policyholder for 100% of health care costs between 20 and 120 less the deductible. Health care costs above 120 are reimbursed at 50%. Let G be the cumulative distribution function of reimbursements given that the reimbursement is positive. Calculate $G(115)$.

37 Gamma Distribution

We start this section by introducing the **Gamma function** defined by

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy, \quad \alpha > 0.$$

For example,

$$\Gamma(1) = \int_0^\infty e^{-y} dy = -e^{-y} \,|_0^\infty = 1.$$

For $\alpha > 1$ we can use integration by parts with $u = y^{\alpha-1}$ and $dv = e^{-y}dy$ to obtain

$$\begin{aligned}
\Gamma(\alpha) =& - e^{-y} y^{\alpha-1} \,|_0^\infty + \int_0^\infty e^{-y}(\alpha-1)y^{\alpha-2}dy \\
=&(\alpha-1)\int_0^\infty e^{-y} y^{\alpha-2} dy \\
=&(\alpha-1)\Gamma(\alpha-1)
\end{aligned}$$

If n is a positive integer greater than 1 then by applying the previous relation repeatedly we find

$$\begin{aligned}
\Gamma(n) =&(n-1)\Gamma(n-1) \\
=&(n-1)(n-2)\Gamma(n-2) \\
&\vdots \\
=&(n-1)(n-2)\cdots 3\cdot 2\cdot \Gamma(1) = (n-1)!
\end{aligned}$$

Example 37.1
Show that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Solution.
Using the substitution $y = \frac{z^2}{2}$, we find

$$\begin{aligned}
\Gamma\left(\frac{1}{2}\right) =& \int_0^\infty y^{-\frac{1}{2}} e^{-y} dy = \sqrt{2}\int_0^\infty e^{-\frac{z^2}{2}} dz \\
=&\frac{\sqrt{2}}{2}\sqrt{2\pi}\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{z^2}{2}} dz\right] \\
=&\sqrt{\pi}
\end{aligned}$$

where we used the fact that Z is the standard normal distribution with

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}dz = 1 \blacksquare$$

A **Gamma** random variable with parameters $\alpha > 0$ and $\lambda > 0$ has a pdf

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

We call α the **shape parameter** because changing α changes the shape of the density function. We call λ the **scale parameter** because if X is a gamma distribution with parameters (α, λ) then cX is also a gamma distribution with parameters $(\alpha, \frac{\lambda}{c})$ where $c > 0$ is a constant. See Problem 37.1. The parameter λ rescales the density function without changing its shape.

To see that $f(t)$ is indeed a probability density function we have

$$\begin{aligned} \Gamma(\alpha) &= \int_0^{\infty} e^{-x}x^{\alpha-1}dx \\ 1 &= \int_0^{\infty} \frac{e^{-x}x^{\alpha-1}}{\Gamma(\alpha)}dx \\ 1 &= \int_0^{\infty} \frac{\lambda e^{-\lambda y}(\lambda y)^{\alpha-1}}{\Gamma(\alpha)}dy \end{aligned}$$

where we used the substitution $x = \lambda y$.
The gamma distribution is skewed right as shown in Figure 37.1

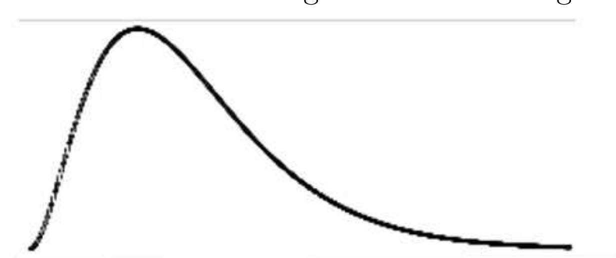


Figure 37.1

Note that the above computation involves a $\Gamma(\alpha)$ integral. Thus, the origin of the name of the random variable.

The cdf of the gamma distribution is

$$F(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^x y^{\alpha-1} e^{-\lambda y} dy.$$

The following reduction formula is useful when computing $F(x)$:

$$\int x^n e^{-\lambda x} dx = -\frac{1}{\lambda} x^n e^{-\lambda x} + \frac{n}{\lambda} \int x^{n-1} e^{-\lambda x} dx. \tag{37.1}$$

Example 37.2
Let X be a gamma random variable with $\alpha = 4$ and $\lambda = \frac{1}{2}$. Compute $\Pr(2 < X < 4)$.

Solution.
We have

$$\begin{aligned} \Pr(2 < X < 4) =& \int_2^4 \frac{1}{2^4 \Gamma(4)} x^3 e^{-\frac{x}{2}} dx \\ =& \frac{1}{96} \int_2^4 x^3 e^{-\frac{x}{2}} dx \approx 0.124 \end{aligned}$$

where we used the reduction formula (37.1) ■

The next result provides formulas for the expected value and the variance of a gamma distribution.

Theorem 37.1
If X is a Gamma random variable with parameters (λ, α) then
(a) $E(X) = \frac{\alpha}{\lambda}$
(b) $Var(X) = \frac{\alpha}{\lambda^2}$.

Solution.
(a)

$$\begin{aligned} E(X) =& \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda x e^{-\lambda x} (\lambda x)^{\alpha-1} dx \\ =& \frac{1}{\lambda \Gamma(\alpha)} \int_0^\infty \lambda e^{-\lambda x} (\lambda x)^{\alpha} dx \\ =& \frac{\Gamma(\alpha+1)}{\lambda \Gamma(\alpha)} \\ =& \frac{\alpha}{\lambda} \end{aligned}$$

(b)

$$\begin{aligned}
E(X^2) =&\frac{1}{\Gamma(\alpha)}\int_0^\infty x^2e^{-\lambda x}\lambda^\alpha x^{\alpha-1}dx\\
=&\frac{1}{\Gamma(\alpha)}\int_0^\infty x^{\alpha+1}\lambda^\alpha e^{-\lambda x}dx\\
=&\frac{\Gamma(\alpha+2)}{\lambda^2\Gamma(\alpha)}\int_0^\infty \frac{x^{\alpha+1}\lambda^{\alpha+2}e^{-\lambda x}}{\Gamma(\alpha+2)}dx\\
=&\frac{\Gamma(\alpha+2)}{\lambda^2\Gamma(\alpha)}
\end{aligned}$$

where the last integral is the integral of the pdf of a Gamma random variable with parameters $(\alpha+2,\lambda)$. Thus,

$$E(X^2)=\frac{\Gamma(\alpha+2)}{\lambda^2\Gamma(\alpha)}=\frac{(\alpha+1)\Gamma(\alpha+1)}{\lambda^2\Gamma(\alpha)}=\frac{\alpha(\alpha+1)}{\lambda^2}.$$

Finally,

$$Var(X)=E(X^2)-(E(X))^2=\frac{\alpha(\alpha+1)}{\lambda^2}-\frac{\alpha^2}{\lambda^2}=\frac{\alpha}{\lambda^2} \blacksquare$$

Example 37.3
In a certain city, the daily consumption of water (in millions of liters) can be treated as a random variable having a gamma distribution with $\alpha=3$ and $\lambda=0.5$.
(a) What is the random variable? What is the expected daily consumption?
(b) If the daily capacity of the city is 12 million liters, what is the probability that this water supply will be inadequate on a given day? Set up the appropriate integral but do not evaluate.
(c) What is the variance of the daily consumption of water?

Solution.
(a) The random variable is the daily consumption of water in millions of liters. The expected daily consumption is the expected value of a gamma distributed variable with parameters $\alpha=3$ and $\lambda=\frac{1}{2}$ which is $E(X)=\frac{\alpha}{\lambda}=6$.
(b) The probability is $\frac{1}{2^3\Gamma(3)}\int_{12}^\infty x^2e^{-\frac{x}{2}}dx=\frac{1}{16}\int_{12}^\infty x^2e^{-\frac{x}{2}}dx$.
(c) The variance is

$$\text{Var}(X)=\frac{3}{0.5^2}=12 \blacksquare$$

It is easy to see that when the parameter set is restricted to $(\alpha, \lambda) = (1, \lambda)$ the gamma distribution becomes the exponential distribution. Another interesting special case is when the parameter set is $(\alpha, \lambda) = \left(\frac{n}{2}, \frac{1}{2}\right)$ where n is a positive integer. This distribution is called the **chi-squared** distribution with **degrees of freedom** n.. The chi-squared random variable is usually denoted by χ_n^2.

The gamma random variable can be used to model the waiting time required for α events to occur, given that the events occur randomly in a Poisson process with mean time between events equals to λ^{-1}.

Example 37.4
On average, it takes you 35 minutes to hunt a duck. Suppose that you want to bring home exactly 3 ducks. What is the probability you will need between 1 and 2 hours to hunt them?

Solution.
Let X be the time in minutes to hunt the 3 ducks. Then X is a gamma random variable with $\lambda = \frac{1}{35}$ duck per minute and $\alpha = 3$ ducks. Thus, $\Pr(60 < X < 120) = \int_{60}^{120} \frac{1}{85750} e^{-\frac{x}{35}} x^2 dx \approx 0.419$ where we used (37.1) ■

Practice Problems

Problem 37.1
Let X be a gamma distribution with parameters (α, λ). Let $Y = cX$ with $c > 0$. Show that
$$F_Y(y) = \frac{(\lambda/c)^\alpha}{\Gamma(\alpha)} \int_0^y z^{\alpha-1} e^{-\lambda\frac{z}{c}} dz.$$
Hence, Y is a gamma distribution with parameters $\left(\alpha, \frac{\lambda}{c}\right)$.

Problem 37.2
If X has a probability density function given by

$$f(x) = \begin{cases} 4x^2 e^{-2x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find the mean and the variance.

Problem 37.3
Let X be a gamma random variable with $\lambda = 1.8$ and $\alpha = 3$. Compute $\Pr(X > 3)$.

Problem 37.4
Suppose the time (in hours) taken by a technician to fix a computer is a random variable X having a gamma distribution with parameters $\alpha = 3$ and $\lambda = 0.5$. What is the probability that it takes at most 1 hour to fix a computer?

Problem 37.5
Suppose the continuous random variable X has the following pdf:

$$f(x) = \begin{cases} \frac{1}{16} x^2 e^{-\frac{x}{2}} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find $E(X^3)$.

Problem 37.6
Let X be the standard normal distribution. Show that X^2 is a gamma distribution with $\alpha = \lambda = \frac{1}{2}$.

Problem 37.7
Let X be a gamma random variable with parameter (α, λ). Find $E(e^{tX})$.

Problem 37.8
Show that the gamma density function with parameters $\alpha > 1$ and $\lambda > 0$ has a relative maximum at $x = \frac{1}{\lambda}(\alpha - 1)$.

Problem 37.9
Let X be a gamma distribution with parameters $\alpha = 3$, and $\lambda = \frac{1}{6}$.
(a) Give the density function, as well as the mean and standard deviation of X.
(b) Find $E(3X^2 + X - 1)$.

Problem 37.10
Find the pdf, mean and variance of the chi-squared distribution with dgrees of freedom n.

38 The Distribution of a Function of a Random Variable

Let X be a continuous random variable. Let $g(x)$ be a function. Then $g(X)$ is also a random variable. In this section we are interested in finding the probability density function of $g(X)$.
The following example illustrates the method of finding the probability density function by finding first its cdf.

Example 38.1
If the probability density of X is given by

$$f(x) = \begin{cases} 6x(1-x) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

find the probability density of $Y = X^3$.

Solution.
We have

$$\begin{aligned} F(y) =& \Pr(Y \le y) = \Pr(X^3 \le y) = \Pr(X \le y^{\frac{1}{3}}) \\ =& \int_0^{y^{\frac{1}{3}}} 6x(1-x)dx = 3y^{\frac{2}{3}} - 2y \end{aligned}$$

Hence, $f(y) = F'(y) = 2(y^{-\frac{1}{3}} - 1)$, for $0 < y < 1$ and 0 otherwise ■

Example 38.2
Let X be a random variable with probability density $f(x)$. Find the probability density function of $Y = |X|$.

Solution.
Clearly, $F_Y(y) = 0$ for $y \le 0$. So assume that $y > 0$. Then

$$\begin{aligned} F_Y(y) =& \Pr(Y \le y) = \Pr(|X| \le y) \\ =& \Pr(-y \le X \le y) = F_X(y) - F_X(-y) \end{aligned}$$

Thus, $f_Y(y) = F'_Y(y) = f_X(y) + f_X(-y)$ for $y > 0$ and 0 otherwise ■

The following theorem provides a formula for finding the probability density of $g(X)$ for monotone g without the need for finding the distribution function.

Theorem 38.1
Let X be a continuous random variable with pdf f_X. Let $g(x)$ be a monotone and differentiable function of x. Suppose that $g^{-1}(Y) = X$. Then the random variable $Y = g(X)$ has a pdf given by

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

Proof.
Suppose first that $g(\cdot)$ is increasing. Then

$$\begin{aligned}F_Y(y) =&\Pr(Y \leq y) = \Pr(g(X) \leq y)\\ =&\Pr(X \leq g^{-1}(y)) = F_X(g^{-1}(y))\end{aligned}$$

Differentiating and using the chain rule, we find

$$f_Y(y) = \frac{dF_Y(y)}{dy} = f_X[g^{-1}(y)]\frac{d}{dy}g^{-1}(y).$$

Now, suppose that $g(\cdot)$ is decreasing. Then

$$\begin{aligned}F_Y(y) =&\Pr(Y \leq y) = \Pr(g(X) \leq y)\\ =&\Pr(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y))\end{aligned}$$

Differentiating we find

$$f_Y(y) = \frac{dF_Y(y)}{dy} = -f_X[g^{-1}(y)]\frac{d}{dy}g^{-1}(y) \blacksquare$$

Example 38.3
Let X be a continuous random variable with pdf f_X. Find the pdf of $Y = -X$.

Solution.
By the previous theorem we have

$$f_Y(y) = f_X(-y) \blacksquare$$

Example 38.4
Let X be a continuous random variable with pdf f_X. Find the pdf of $Y = aX + b,\ a > 0$.

Solution.
Let $g(x) = ax + b$. Then $g^{-1}(y) = \frac{y-b}{a}$. By the previous theorem, we have

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right) \blacksquare$$

Example 38.5
Suppose X is a random variable with the following density :

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

(a) Find the cdf of $|X|$.
(b) Find the pdf of X^2.

Solution.
(a) $|X|$ takes values in $[0, \infty)$. Thus, $F_{|X|}(x) = 0$ for $x \le 0$. Now, for $x > 0$ we have

$$F_{|X|}(x) = \Pr(|X| \le x) = \int_{-x}^{x} \frac{1}{\pi(x^2+1)} dx = \frac{2}{\pi} \tan^{-1} x.$$

Hence,

$$F_{|X|}(x) = \begin{cases} 0 & x \le 0 \\ \frac{2}{\pi} \tan^{-1} x & x > 0. \end{cases}$$

(b) X^2 also takes only nonnegative values, so the density $f_{X^2}(x) = 0$ for $x \le 0$. Furthermore, $F_{X^2}(x) = \Pr(X^2 \le x) = \Pr(|X| \le \sqrt{x}) = \frac{2}{\pi} \tan^{-1} \sqrt{x}$. So by differentiating we get

$$f_{X^2}(x) = \begin{cases} 0 & x \le 0 \\ \frac{1}{\pi\sqrt{x}(1+x)} & x > 0 \end{cases} \blacksquare$$

Remark 38.1
In general, if a function does not have a unique inverse, we must sum over all possible inverse values.

Example 38.6
Let X be a continuous random variable with pdf f_X. Find the pdf of $Y = X^2$.

Solution.
Let $g(x) = x^2$. Then $g^{-1}(y) = \pm\sqrt{y}$. Thus,

$$F_Y(y) = \Pr(Y \le y) = \Pr(X^2 \le y) = \Pr(-\sqrt{y} \le X \le \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Differentiate both sides to obtain,

$$f_Y(y) = \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}} \quad \blacksquare$$

Practice Problems

Problem 38.1
Suppose $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2}}$ and let $Y = aX + b$. Find $f_Y(y)$.

Problem 38.2
Let X be a continuous random variable with pdf

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find probability density function for $Y = 3X - 1$.

Problem 38.3
Let X be a random variable with density function

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the density function of $Y = 8X^3$.

Problem 38.4
Suppose X is an exponential random variable with density function

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

What is the density function of $Y = e^X$?

Problem 38.5
Gas molecules move about with varying velocity which has, according to the Maxwell- Boltzmann law, a probability density given by

$$f(v) = cv^2 e^{-\beta v^2}, \quad v \geq 0$$

The kinetic energy is given by $Y = E = \frac{1}{2}mv^2$ where m is the mass. What is the density function of Y?

Problem 38.6
Let X be a random variable that is uniformly distributed in (0,1). Find the probability density function of $Y = -\ln X$.

Problem 38.7
Let X be a uniformly distributed function over $[-\pi, \pi]$. That is

$$f(x) = \begin{cases} \frac{1}{2\pi} & -\pi \leq x \leq \pi \\ 0 & \text{otherwise} \end{cases}$$

Find the probability density function of $Y = \cos X$.

Problem 38.8
Suppose X has the uniform distribution on (0, 1). Compute the probability density function and expected value of:
(a) $X^\alpha, \quad \alpha > 0$ (b) $\ln X$ (c) e^X (d) $\sin \pi X$

Problem 38.9 ‡
The time, T, that a manufacturing system is out of operation has cumulative distribution function

$$F(t) = \begin{cases} 1 - \left(\frac{2}{t}\right)^2 & t > 2 \\ 0 & \text{otherwise} \end{cases}$$

The resulting cost to the company is $Y = T^2$. Determine the density function of Y, for $y > 4$.

Problem 38.10 ‡
An investment account earns an annual interest rate R that follows a uniform distribution on the interval $(0.04, 0.08)$. The value of a 10,000 initial investment in this account after one year is given by $V = 10,000e^R$.
Determine the cumulative distribution function, $F_V(v)$ of V.

Problem 38.11 ‡
An actuary models the lifetime of a device using the random variable $Y = 10X^{0.8}$, where X is an exponential random variable with mean 1 year.
Determine the probability density function $f_Y(y)$, for $y > 0$, of the random variable Y.

Problem 38.12 ‡
Let T denote the time in minutes for a customer service representative to respond to 10 telephone inquiries. T is uniformly distributed on the interval with endpoints 8 minutes and 12 minutes. Let R denote the average rate, in customers per minute, at which the representative responds to inquiries.
Find the density function $f_R(r)$ of R.

Problem 38.13 ‡
The monthly profit of Company A can be modeled by a continuous random variable with density function f_A. Company B has a monthly profit that is twice that of Company A.
Determine the probability density function of the monthly profit of Company B.

Problem 38.14
Let X have normal distribution with mean 1 and standard deviation 2.
(a) Find $\Pr(|X| \leq 1)$.
(b) Let $Y = e^X$. Find the probability density function $f_Y(y)$ of Y.

Problem 38.15
Let X be a uniformly distributed random variable on the interval $(-1, 1)$. Show that $Y = X^2$ is a beta random variable with paramters $(\frac{1}{2}, 1)$.

Problem 38.16
Let X be a random variable with density function

$$f(x) = \begin{cases} \frac{3}{2}x^2 & -1 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the pdf of $Y = 3X$.
(b) Find the pdf of $Z = 3 - X$.

Problem 38.17
Let X be a continuous random variable with density function

$$f(x) = \begin{cases} 1 - |x| & -1 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find the density function of $Y = X^2$.

Problem 38.18
If $f(x) = xe^{-\frac{x^2}{2}}$, for $x > 0$ and $Y = \ln X$, find the density function for Y.

Problem 38.19
Let X be a continuous random variable with pdf

$$f(x) = \begin{cases} 2(1 - x) & 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the pdf of $Y = 10X - 2$.
(b) Find the expected value of Y.
(c) Find $\Pr(Y < 0)$.

Joint Distributions

There are many situations which involve the presence of several random variables and we are interested in their joint behavior. This chapter is concerned with the joint probability structure of two or more random variables defined on the same sample space.

39 Jointly Distributed Random Variables

Suppose that X and Y are two random variables defined on the same sample space S. The **joint cumulative distribution function** of X and Y is the function

$$F_{XY}(x,y) = \Pr(X \leq x, Y \leq y) = \Pr(\{e \in S : X(e) \leq x \text{ } and \text{ } Y(e) \leq y\}).$$

Example 39.1
Consider the experiment of throwing a fair coin and a fair die simultaneously. The sample space is

$$S = \{(H,1),(H,2),\cdots,(H,6),(T,1),(T,2),\cdots,(T,6)\}.$$

Let X be the number of head showing on the coin, $X \in \{0,1\}$. Let Y be the number showing on the die, $Y \in \{1,2,3,4,5,6\}$. Thus, if $e = (H,1)$ then $X(e) = 1$ and $Y(e) = 1$. Find $F_{XY}(1,2)$.

Solution.

$$\begin{aligned}F_{XY}(1,2) =&\Pr(X \leq 1, Y \leq 2)\\ =&\Pr(\{(H,1),(H,2),(T,1),(T,2)\})\\ =&\frac{4}{12} = \frac{1}{3} \blacksquare\end{aligned}$$

In what follows, individual cdfs will be referred to as **marginal distributions**. These cdfs are obtained from the joint cumulative distribution as follows

$$\begin{aligned}
F_X(x) =&\Pr(X \le x)\\
=&\Pr(X \le x, Y < \infty)\\
=&\Pr(\lim_{y\to\infty}\{X \le x, Y \le y\})\\
=&\lim_{y\to\infty}\Pr(X \le x, Y \le y)\\
=&\lim_{y\to\infty}F_{XY}(x,y) = F_{XY}(x,\infty).
\end{aligned}$$

In a similar way, one can show that

$$F_Y(y) = \lim_{x\to\infty}F_{XY}(x,y) = F_{XY}(\infty,y).$$

It is easy to see that

$$F_{XY}(\infty,\infty) = \Pr(X < \infty, Y < \infty) = 1.$$

Also,

$$F_{XY}(-\infty,y) = 0.$$

This follows from

$$0 \le F_{XY}(-\infty,y) = \Pr(X < -\infty, Y \le y) \le \Pr(X < -\infty) = F_X(-\infty) = 0.$$

Similarly,

$$F_{XY}(x,-\infty) = 0.$$

All joint probability statements about X and Y can be answered in terms of their joint distribution functions. For example,

$$\begin{aligned}
\Pr(X > x, Y > y) =&1 - \Pr(\{X > x, Y > y\}^c)\\
=&1 - \Pr(\{X > x\}^c \cup \{Y > y\}^c)\\
=&1 - [\Pr(\{X \le x\} \cup \{Y \le y\})\\
=&1 - [\Pr(X \le x) + \Pr(Y \le y) - \Pr(X \le x, Y \le y)]\\
=&1 - F_X(x) - F_Y(y) + F_{XY}(x,y).
\end{aligned}$$

Also, if $a_1 < a_2$ and $b_1 < b_2$ then

$$\begin{aligned}\Pr(a_1 < X \leq a_2, b_1 < Y \leq b_2) =& \Pr(X \leq a_2, Y \leq b_2) - \Pr(X \leq a_2, Y \leq b_1) \\ & - \Pr(X \leq a_1, Y \leq b_2) + \Pr(X \leq a_1, Y \leq b_1) \\ =& F_{XY}(a_2, b_2) - F_{XY}(a_1, b_2) - F_{XY}(a_2, b_1) + F_{XY}(a_1, b_1).\end{aligned}$$

This is clear if you use the concept of area shown in Figure 39.1

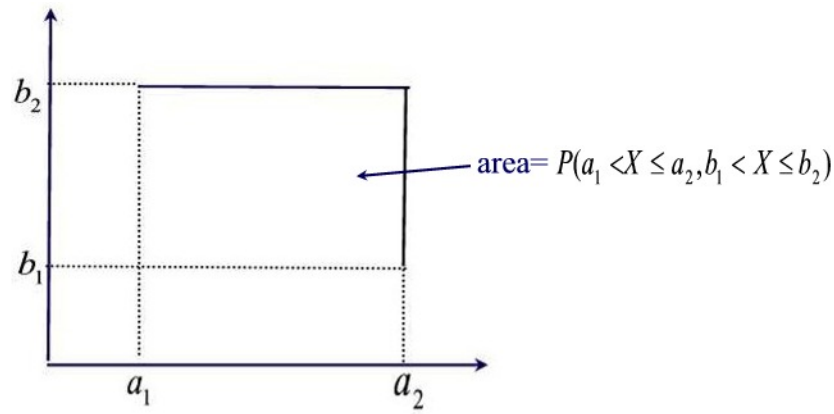


Figure 39.1

If X and Y are both discrete random variables, we define the **joint probability mass function** of X and Y by

$$p_{XY}(x, y) = \Pr(X = x, Y = y).$$

The marginal probability mass function of X can be obtained from $p_{XY}(x, y)$ by

$$p_X(x) = \Pr(X = x) = \sum_{y: p_{XY}(x,y) > 0} p_{XY}(x, y).$$

Similarly, we can obtain the marginal pmf of Y by

$$p_Y(y) = \Pr(Y = y) = \sum_{x: p_{XY}(x,y) > 0} p_{XY}(x, y).$$

This simply means that to find the probability that X takes on a specific value we sum across the row associated with that value. To find the probability that Y takes on a specific value we sum the column associated with that value as illustrated in the next example.

Example 39.2
A fair coin is tossed 4 times. Let the random variable X denote the number of heads in the first 3 tosses, and let the random variable Y denote the number of heads in the last 3 tosses.
(a) What is the joint pmf of X and Y?
(b) What is the probability 2 or 3 heads appear in the first 3 tosses and 1 or 2 heads appear in the last three tosses?
(c) What is the joint cdf of X and Y?
(d) What is the probability less than 3 heads occur in both the first and last 3 tosses?
(e) Find the probability that one head appears in the first three tosses.

Solution.
(a) The joint pmf is given by the following table

$X\backslash Y$	0	1	2	3	$p_X(.)$
0	1/16	1/16	0	0	2/16
1	1/16	3/16	2/16	0	6/16
2	0	2/16	3/16	1/16	6/16
3	0	0	1/16	1/16	2/16
$p_Y(.)$	2/16	6/16	6/16	2/16	1

(b) $\Pr((X,Y) \in \{(2,1),(2,2),(3,1),(3,2)\}) = \Pr(2,1)+\Pr(2,2)+\Pr(3,1)+\Pr(3,2) = \frac{3}{8}$
(c) The joint cdf is given by the following table

$X\backslash Y$	0	1	2	3
0	1/16	2/16	2/16	2/16
1	2/16	6/16	8/16	8/16
2	2/16	8/16	13/16	14/16
3	2/16	8/16	14/16	1

(d) $\Pr(X<3, Y<3) = F(2,2) = \frac{13}{16}$
(e) $\Pr(X=1) = \Pr((X,Y) \in \{(1,0),(1,1),(1,2),(1,3)\}) = 1/16 + 3/16 + 2/16 = \frac{3}{8}$ ■

Example 39.3
Suppose two balls are chosen from a box containing 3 white, 2 red and 5 blue balls. Let X = the number of white balls chosen and Y = the number of blue balls chosen. Find the joint pmf of X and Y.

Solution.

$$\begin{aligned}
p_{XY}(0,0) =&\frac{{}_2C_2}{{}_{10}C_2} = \frac{1}{45}\\
p_{XY}(0,1) =&\frac{{}_2C_1 \cdot {}_5C_1}{{}_{10}C_2} = \frac{10}{45}\\
p_{XY}(0,2) =&\frac{{}_5C_2}{{}_{10}C_2} = \frac{10}{45}\\
p_{XY}(1,0) =&\frac{{}_2C_1 \cdot {}_3C_1}{{}_{10}C_2} = \frac{6}{45}\\
p_{XY}(1,1) =&\frac{{}_5C_1 \cdot {}_3C_1}{{}_{10}C_2} = \frac{15}{45}\\
p_{XY}(1,2) =&0\\
p_{XY}(2,0) =&\frac{{}_3C_2}{{}_{10}C_2} = \frac{3}{45}\\
p_{XY}(2,1) =&0\\
p_{XY}(2,2) =&0
\end{aligned}$$

The pmf of X is

$$\begin{aligned}
p_X(0) =&\Pr(X=0) = \sum_{y:p_{XY}(0,y)>0} p_{XY}(0,y) = \frac{1+10+10}{45} = \frac{21}{45}\\
p_X(1) =&\Pr(X=1) = \sum_{y:p_{XY}(1,y)>0} p_{XY}(1,y) = \frac{6+15}{45} = \frac{21}{45}\\
p_X(2) =&\Pr(X=2) = \sum_{y:p_{XY}(2,y)>0} p_{XY}(2,y) = \frac{3}{45} = \frac{3}{45}
\end{aligned}$$

The pmf of y is

$$\begin{aligned}
p_Y(0) =&\Pr(Y=0) = \sum_{x:p_{XY}(x,0)>0} p_{XY}(x,0) = \frac{1+6+3}{45} = \frac{10}{45}\\
p_Y(1) =&\Pr(Y=1) = \sum_{x:p_{XY}(x,1)>0} p_{XY}(x,1) = \frac{10+15}{45} = \frac{25}{45}\\
p_Y(2) =&\Pr(Y=2) = \sum_{x:p_{XY}(x,2)>0} p_{XY}(x,2) = \frac{10}{45} = \frac{10}{45} \blacksquare
\end{aligned}$$

Two random variables X and Y are said to be **jointly continuous** if there exists a function $f_{XY}(x,y) \geq 0$ with the property that for every subset C of $\mathbb{R}^2$ we have

$$\Pr((X,Y) \in C) = \iint_{(x,y)\in C} f_{XY}(x,y)dxdy$$

The function $f_{XY}(x,y)$ is called the **joint probability density function** of X and Y.
If A and B are any sets of real numbers then by letting $C = \{(x,y) : x \in A, y \in B\}$ we have

$$\Pr(X \in A, Y \in B) = \int_B \int_A f_{XY}(x,y)dxdy$$

As a result of this last equation we can write

$$\begin{aligned} F_{XY}(x,y) =& \Pr(X \in (-\infty, x], Y \in (-\infty, y]) \\ =& \int_{-\infty}^{y} \int_{-\infty}^{x} f_{XY}(u,v)dudv \end{aligned}$$

It follows upon differentiation that

$$f_{XY}(x,y) = \frac{\partial^2}{\partial y \partial x} F_{XY}(x,y)$$

whenever the partial derivatives exist.

Example 39.4
The cumulative distribution function for the joint distribution of the continuous random variables X and Y is $F_{XY}(x,y) = 0.2(3x^3y + 2x^2y^2),\ 0 \leq x \leq 1, 0 \leq y \leq 1$. Find $f_{XY}(\frac{1}{2}, \frac{1}{2})$.

Solution.
Since

$$f_{XY}(x,y) = \frac{\partial^2}{\partial y \partial x} F_{XY}(x,y) = 0.2(9x^2 + 8xy)$$

we find $f_{XY}(\frac{1}{2}, \frac{1}{2}) = \frac{17}{20}$ ■

Now, if X and Y are jointly continuous then they are individually continuous, and their probability density functions can be obtained as follows:

$$\begin{aligned}\Pr(X \in A) =& \Pr(X \in A, Y \in (-\infty, \infty))\\ =& \int_A \int_{-\infty}^{\infty} f_{XY}(x,y)dydx\\ =& \int_A f_X(x)dx\end{aligned}$$

where

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy$$

is thus the probability density function of X. Similarly, the probability density function of Y is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx.$$

Example 39.5
Let X and Y be random variables with joint pdf

$$f_{XY}(x,y) = \begin{cases} \frac{1}{4} & -1 \le x, y \le 1\\ 0 & \text{Otherwise}\end{cases}$$

Determine
(a) $\Pr(X^2 + Y^2 < 1)$,
(b) $\Pr(2X - Y > 0)$,
(c) $\Pr(|X + Y| < 2)$.

Solution.
(a)

$$\Pr(X^2 + Y^2 < 1) = \int_0^{2\pi} \int_0^1 \frac{1}{4} r dr d\theta = \frac{\pi}{4}.$$

(b)

$$\Pr(2X - Y > 0) = \int_{-1}^{1} \int_{\frac{y}{2}}^{1} \frac{1}{4} dx dy = \frac{1}{2}.$$

Note that $\Pr(2X - Y > 0)$ is the area of the region bounded by the lines $y = 2x, x = -1, x = 1, y = -1$ and $y = 1$. A graph of this region will help

you understand the integration process used above.
(c) Since the square with vertices $(1,1),(1,-1),(-1,1),(-1,-1)$ is completely contained in the region $-2 < x+y < 2$, we have

$$\Pr(|X+Y| < 2) = 1 \blacksquare$$

Remark 39.1
Joint pdfs and joint cdfs for three or more random variables are obtained as straightforward generalizations of the above definitions and conditions.

Practice Problems

Problem 39.1

A security check at an airport has two express lines. Let X and Y denote the number of customers in the first and second line at any given time. The joint probability function of X and $Y, p_{XY}(x,y)$, is summarized by the following table

$X\backslash Y$	0	1	2	3	$p_X(.)$
0	0.1	0.2	0	0	0.3
1	0.2	0.25	0.05	0	0.5
2	0	0.05	0.05	0.025	0.125
3	0	0	0.025	0.05	0.075
$p_Y(.)$	0.3	0.5	0.125	0.075	1

(a) Show that $p_{XY}(x,y)$ is a joint probability mass function.
(b) Find the probability that more than two customers are in line.
(c) Find $\Pr(|X-Y|=1)$.
(d) Find $p_X(x)$.

Problem 39.2

Given:

$X\backslash Y$	1	2	3	$p_X(.)$
1	0.1	0.05	0.02	0.17
2	0.1	0.35	0.05	0.50
3	0.03	0.1	0.2	0.33
$p_Y(.)$	0.23	0.50	0.27	1

Find $\Pr(X \geq 2, Y \geq 3)$.

Problem 39.3

Given:

$X\backslash Y$	0	1	2	$p_X(.)$
0	0.4	0.12	0.08	0.6
1	0.15	0.08	0.03	0.26
2	0.1	0.03	0.01	0.14
$p_Y(.)$	0.65	0.23	0.12	1

Find the following: (a) $\Pr(X = 0, Y = 2)$.
(b) $\Pr(X > 0, Y \leq 1)$.
(c) $\Pr(X \leq 1)$.
(d) $\Pr(Y > 0)$.
(e) $\Pr(X = 0)$.
(f) $\Pr(Y = 0)$.
(g) $\Pr(X = 0, Y = 0)$.

Problem 39.4
Given:

$X\backslash Y$	15	16	$p_X(.)$
129	0.12	0.08	0.2
130	0.4	0.30	0.7
131	0.06	0.04	0.1
$p_Y(.)$	0.58	0.42	1

(a) Find $\Pr(X = 130, Y = 15)$.
(b) Find $\Pr(X \geq 130, Y \geq 15)$.

Problem 39.5
Suppose the random variables X and Y have a joint pdf

$$f_{XY}(x,y) = \begin{cases} \frac{20-x-y}{375} & 0 \leq x, y \leq 5 \\ 0 & \text{otherwise.} \end{cases}$$

Find $\Pr(1 \leq X \leq 2, 2 \leq Y \leq 3)$.

Problem 39.6
Assume the joint pdf of X and Y is

$$f_{XY}(x,y) = \begin{cases} xye^{-\frac{x^2+y^2}{2}} & 0 < x, y \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $F_{XY}(x, y)$.
(b) Find $f_X(x)$ and $f_Y(y)$.

Problem 39.7
Show that the following function is not a joint probability density function?

$$f_{XY}(x,y) = \begin{cases} x^a y^{1-a} & 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

where $0 < a < 1$. What factor should you multiply $f_{XY}(x,y)$ to make it a joint probability density function?

Problem 39.8 ‡
A device runs until either of two components fails, at which point the device stops running. The joint density function of the lifetimes of the two components, both measured in hours, is

$$f_{XY}(x,y) = \begin{cases} \frac{x+y}{8} & 0 < x, y < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the probability that the device fails during its first hour of operation?

Problem 39.9 ‡
An insurance company insures a large number of drivers. Let X be the random variable representing the company's losses under collision insurance, and let Y represent the company's losses under liability insurance. X and Y have joint density function

$$f_{XY}(x,y) = \begin{cases} \frac{2x+2-y}{4} & 0 < x < 1, 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the probability that the total loss is at least 1 ?

Problem 39.10 ‡
A car dealership sells 0, 1, or 2 luxury cars on any day. When selling a car, the dealer also tries to persuade the customer to buy an extended warranty for the car. Let X denote the number of luxury cars sold in a given day, and let Y denote the number of extended warranties sold. Given the following information

$$\begin{aligned}
\Pr(X=0,Y=0) =& \frac{1}{6} \\
\Pr(X=1,Y=0) =& \frac{1}{12} \\
\Pr(X=1,Y=1) =& \frac{1}{6} \\
\Pr(X=2,Y=0) =& \frac{1}{12} \\
\Pr(X=2,Y=1) =& \frac{1}{3} \\
\Pr(X=2,Y=2) =& \frac{1}{6}
\end{aligned}$$

What is the variance of X?

Problem 39.11 ‡
A company is reviewing tornado damage claims under a farm insurance policy. Let X be the portion of a claim representing damage to the house and let Y be the portion of the same claim representing damage to the rest of the property. The joint density function of X and Y is

$$f_{XY}(x,y) = \begin{cases} 6[1-(x+y)] & x>0, y>0, x+y<1 \\ 0 & \text{otherwise.} \end{cases}$$

Determine the probability that the portion of a claim representing damage to the house is less than 0.2.

Problem 39.12 ‡
Let X and Y be continuous random variables with joint density function

$$f_{XY}(x,y) = \begin{cases} 15y & x^2 \le y \le x \\ 0 & \text{otherwise.} \end{cases}$$

Find the marginal density function of Y.

Problem 39.13 ‡
Let X represent the age of an insured automobile involved in an accident. Let Y represent the length of time the owner has insured the automobile at the time of the accident. X and Y have joint probability density function

$$f_{XY}(x,y) = \begin{cases} \frac{1}{64}(10-xy^2) & 2 \le x \le 10, 0 \le y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the expected age of an insured automobile involved in an accident.

Problem 39.14 ‡
A device contains two circuits. The second circuit is a backup for the first, so the second is used only when the first has failed. The device fails when and only when the second circuit fails.
Let X and Y be the times at which the first and second circuits fail, respectively. X and Y have joint probability density function

$$f_{XY}(x,y) = \begin{cases} 6e^{-x}e^{-2y} & 0<x<y<\infty \\ 0 & \text{otherwise.} \end{cases}$$

What is the expected time at which the device fails?

Problem 39.15 ‡
The future lifetimes (in months) of two components of a machine have the following joint density function:

$$f_{XY}(x,y) = \begin{cases} \frac{6}{125000}(50-x-y) & 0 < x < 50-y < 50 \\ 0 & \text{otherwise.} \end{cases}$$

What is the probability that both components are still functioning 20 months from now?

Problem 39.16
Suppose the random variables X and Y have a joint pdf

$$f_{XY}(x,y) = \begin{cases} x+y & 0 \le x,y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find $\Pr(X > \sqrt{Y})$.

Problem 39.17 ‡
Let X and Y be random losses with joint density function

$$f_{XY}(x,y) = e^{-(x+y)}, \quad x > 0, y > 0$$

and 0 otherwise. An insurance policy is written to reimburse $X+Y$. Calculate the probability that the reimbursement is less than 1.

Problem 39.18
Let X and Y be continuous random variables with joint cumulative distribution $F_{XY}(x,y) = \frac{1}{250}(20xy - x^2y - xy^2)$ for $0 \le x \le 5$ and $0 \le y \le 5$. Compute $\Pr(X > 2)$.

Problem 39.19
Let X and Y be continuous random variables with joint density function

$$f_{XY}(x,y) = \begin{cases} xy & 0 \le x \le 2, 0 \le y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find $\Pr(\frac{X}{2} \le Y \le X)$.

Problem 39.20
Let X and Y be random variables with common range $\{1, 2\}$ and such that $\Pr(X = 1) = 0.7, \Pr(X = 2) = 0.3, \Pr(Y = 1) = 0.4, \Pr(Y = 2) = 0.6$, and $\Pr(X = 1, Y = 1) = 0.2$.
(a) Find the joint probability mass function $p_{XY}(x, y)$.
(b) Find the joint cumulative distribution function $F_{XY}(x, y)$.

Problem 39.21 ‡
A device contains two components. The device fails if either component fails. The joint density function of the lifetimes of the components, measured in hours, is $f(s, t)$, where $0 < s < 1$ and $0 < t < 1$.
What is the probability that the device fails during the first half hour of operation?

Problem 39.22 ‡
A client spends X minutes in an insurance agent's waiting room and Y minutes meeting with the agent. The joint density function of X and Y can be modeled by

$$f(x, y) = \begin{cases} \frac{1}{800} e^{\frac{x}{40} + \frac{y}{20}} & \text{for } x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability that a client spends less than 60 minutes at the agent's office. You do NOT have to evaluate the integrals.

40 Independent Random Variables

Let X and Y be two random variables defined on the same sample space S. We say that X and Y are **independent** random variables if and only if for any two sets of real numbers A and B we have

$$\Pr(X \in A, Y \in B) = \Pr(X \in A)\Pr(Y \in B). \tag{40.1}$$

That is, the events $E = \{X \in A\}$ and $F = \{Y \in B\}$ are independent.
The following theorem expresses independence in terms of pdfs.

Theorem 40.1
If X and Y are discrete random variables, then X and Y are independent if and only if

$$p_{XY}(x, y) = p_X(x)p_Y(y)$$

where $p_X(x)$ and $p_Y(y)$ are the marginal pmfs of X and Y respectively. Similar result holds for continuous random variables where sums are replaced by integrals and pmfs are replaced by pdfs.

Proof.
Suppose that X and Y are independent. Then by letting $A = \{x\}$ and $B = \{y\}$ in Equation 40.1 we obtain

$$\Pr(X = x, Y = y) = \Pr(X = x)\Pr(Y = y)$$

that is

$$p_{XY}(x, y) = p_X(x)p_Y(y).$$

Conversely, suppose that $p_{XY}(x, y) = p_X(x)p_Y(y)$. Let A and B be any sets of real numbers. Then

$$\begin{aligned}
\Pr(X \in A, Y \in B) =& \sum_{y \in B} \sum_{x \in A} p_{XY}(x, y) \\
=& \sum_{y \in B} \sum_{x \in A} p_X(x)p_Y(y) \\
=& \sum_{y \in B} p_Y(y) \sum_{x \in A} p_X(x) \\
=& \Pr(Y \in B)\Pr(X \in A)
\end{aligned}$$

and thus equation 40.1 is satisfied. That is, X and Y are independent ■

Example 40.1
A month of the year is chosen at random (each with probability $\frac{1}{12}$). Let X be the number of letters in the month's name, and let Y be the number of days in the month (ignoring leap year).
(a) Write down the joint pdf of X and Y. From this, compute the pdf of X and the pdf of Y.
(b) Find $E(Y)$.
(c) Are the events "$X \leq 6$" and "$Y = 30$" independent?
(d) Are X and Y independent random variables?

Solution.
(a) The joint pdf is given by the following table

Y\ X	3	4	5	6	7	8	9	$p_Y(y)$
28	0	0	0	0	0	$\frac{1}{12}$	0	$\frac{1}{12}$
30	0	$\frac{1}{12}$	$\frac{1}{12}$	0	0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{4}{12}$
31	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	0	$\frac{7}{12}$
$p_X(x)$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{1}{12}$	1

(b) $E(Y) = \left(\frac{1}{12}\right) \times 28 + \left(\frac{4}{12}\right) \times 30 + \left(\frac{7}{12}\right) \times 31 = \frac{365}{12}$
(c) We have $\Pr(X \leq 6) = \frac{6}{12} = \frac{1}{2}$, $\Pr(Y = 30) = \frac{4}{12} = \frac{1}{3}$, $\Pr(X \leq 6, Y = 30) = \frac{2}{12} = \frac{1}{6}$. Since, $\Pr(X \leq 6, Y = 30) = \Pr(X \leq 6)\Pr(Y = 30)$, the two events are independent.
(d) Since $p_{XY}(5, 28) = 0 \neq p_X(5)p_Y(28) = \frac{1}{6} \times \frac{1}{12}$, X and Y are dependent ■

Example 40.2 ‡
Automobile policies are separated into two groups: low-risk and high-risk. Actuary Rahul examines low-risk policies, continuing until a policy with a claim is found and then stopping. Actuary Toby follows the same procedure with high-risk policies. Each low-risk policy has a 10% probability of having a claim. Each high-risk policy has a 20% probability of having a claim. The claim statuses of polices are mutually independent.
Calculate the probability that Actuary Rahul examines fewer policies than Actuary Toby.

Solution.
Let R be the random variable denoting the number of policies examined by Rahul until a claim is found. Then R is a geometric random variable with pmf $p_R(r) = 0.1(0.9)^{r-1}$. Likewise, let T be the random variable denoting the

number of policies examined by Toby until a claim is found. Then T is a geometric random variable with pmf $p_T(t) = 0.2(0.8)^{t-1}$. The joint distribution is given by

$$p_{RT}(r,t) = 0.02(0.9)^{r-1}(0.8)^{t-1}.$$

We want to find $\Pr(R < T)$. We have

$$\begin{aligned}
\Pr(R < T) =& \sum_{r=1}^{\infty}\sum_{t=r+1}^{\infty} 0.02(0.9)^{r-1}(0.8)^{t-1}\\
=& \sum_{r=1}^{\infty} 0.02(0.9)^{r-1}\frac{0.8^r}{1-0.8}\\
=& \frac{0.02}{0.2}\frac{1}{0.9}\sum_{r=1}^{\infty}(0.72)^r\\
=& \frac{1}{9}\frac{0.72}{1-0.72} = 0.2857 \blacksquare
\end{aligned}$$

In the jointly continuous case the condition of independence is equivalent to

$$f_{XY}(x,y) = f_X(x)f_Y(y).$$

It follows from the previous theorem, that if you are given the joint pdf of the random variables X and Y, you can determine whether or not they are independent by calculating the marginal pdfs of X and Y and determining whether or not the relationship $f_{XY}(x,y) = f_X(x)f_Y(y)$ holds.

Example 40.3
The joint pdf of X and Y is given by

$$f_{XY}(x,y) = \begin{cases} 4e^{-2(x+y)} & 0 < x < \infty,\ 0 < y < \infty \\ 0 & \text{Otherwise.} \end{cases}$$

Are X and Y independent?

Solution.
Marginal density $f_X(x)$ is given by

$$f_X(x) = \int_0^{\infty} 4e^{-2(x+y)}dy = 2e^{-2x}\int_0^{\infty} 2e^{-2y}dy = 2e^{-2x},\ x > 0.$$

Similarly, the mariginal density $f_Y(y)$ is given by

$$f_Y(y) = \int_0^\infty 4e^{-2(x+y)}dx = 2e^{-2y}\int_0^\infty 2e^{-2x}dx = 2e^{-2y}, \ y > 0.$$

Now since

$$f_{XY}(x,y) = 4e^{-2(x+y)} = [2e^{-2x}][2e^{-2y}] = f_X(x)f_Y(y)$$

X and Y are independent ■

Example 40.4
The joint pdf of X and Y is given by

$$f_{XY}(x,y) = \begin{cases} 3(x+y) & 0 \le x+y \le 1, \ 0 \le x, y < \infty \\ 0 & \text{Otherwise.} \end{cases}$$

Are X and Y independent?

Solution.
For the limit of integration see Figure 40.1 below.

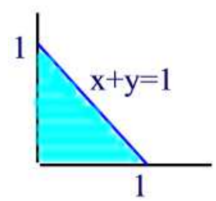


Figure 40.1

The marginal pdf of X is

$$f_X(x) = \int_0^{1-x} 3(x+y)dy = 3xy + \frac{3}{2}y^2\Big|_0^{1-x} = \frac{3}{2}(1-x^2), \ 0 \le x \le 1.$$

The marginal pdf of Y is

$$f_Y(y) = \int_0^{1-y} 3(x+y)dx = \frac{3}{2}x^2 + 3xy\Big|_0^{1-y} = \frac{3}{2}(1-y^2), \ 0 \le y \le 1.$$

But

$$f_{XY}(x,y) = 3(x+y) \ne \frac{3}{2}(1-x^2)\frac{3}{2}(1-y^2) = f_X(x)f_Y(y)$$

so that X and Y are dependent ■

The following theorem provides a necessary and sufficient condition for two random variables to be independent.

Theorem 40.2
Two continuous random variables X and Y are independent if and only if their joint probability density function can be expressed as

$$f_{XY}(x,y) = h(x)g(y), \quad -\infty < x < \infty, -\infty < y < \infty.$$

The same result holds for discrete random variables.

Proof.
Suppose first that X and Y are independent. Then $f_{XY}(x,y) = f_X(x)f_Y(y)$. Let $h(x) = f_X(x)$ and $g(y) = f_Y(y)$.
Conversely, suppose that $f_{XY}(x,y) = h(x)g(y)$. Let $C = \int_{-\infty}^{\infty} h(x)dx$ and $D = \int_{-\infty}^{\infty} g(y)dy$. Then

$$\begin{aligned}CD =& \left(\int_{-\infty}^{\infty} h(x)dx\right)\left(\int_{-\infty}^{\infty} g(y)dy\right)\\ =& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)g(y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1.\end{aligned}$$

Furthermore,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_{-\infty}^{\infty} h(x)g(y)dy = h(x)D$$

and

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx = \int_{-\infty}^{\infty} h(x)g(y)dx = g(y)C.$$

Hence,

$$f_X(x)f_Y(y) = h(x)g(y)CD = h(x)g(y) = f_{XY}(x,y).$$

This proves that X and Y are independent ■

Example 40.5
The joint pdf of X and Y is given by

$$f_{XY}(x,y) = \begin{cases} xye^{-\frac{(x^2+y^2)}{2}} & 0 \le x, y < \infty \\ 0 & \text{Otherwise.}\end{cases}$$

Are X and Y independent?

Solution.
We have

$$f_{XY}(x,y) = xye^{-\frac{(x^2+y^2)}{2}} = xe^{-\frac{x^2}{2}}ye^{-\frac{y^2}{2}}$$

By the previous theorem, X and Y are independent ■

Example 40.6
The joint pdf of X and Y is given by

$$f_{XY}(x,y) = \begin{cases} x+y & 0 \le x, y < 1 \\ 0 & \text{Otherwise.} \end{cases}$$

Are X and Y independent?

Solution.
Let

$$I(x,y) = \begin{cases} 1 & 0 \le x < 1,\ 0 \le y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$f_{XY}(x,y) = (x+y)I(x,y)$$

which clearly does not factor into a part depending only on x and another depending only on y. Thus, by the previous theorem X and Y are dependent ■

Example 40.7 *(Order statistics)*
Let X and Y be two independent random variables with X having a normal distribution with mean μ and variance 1 and Y being the standard normal distribution.
(a) Find the density of $Z = \min\{X, Y\}$.
(b) For each $t \in \mathbb{R}$ calculate $\Pr(\max(X,Y) - \min(X,Y) > t)$.

Solution.
(a) Fix a real number z. Then

$$\begin{aligned} F_Z(z) =& \Pr(Z \le z) = 1 - \Pr(\min(X,Y) > z) \\ =& 1 - \Pr(X > z)\Pr(Y > z) = 1 - (1 - \Phi(z-\mu))(1 - \Phi(z)). \end{aligned}$$

Hence,

$$f_Z(z) = (1 - \Phi(z-\mu))\phi(z) + (1 - \Phi(z))\phi(z-\mu)$$

where $\phi(z)$ is the pdf of the standard normal distribution.
(b) If $t \leq 0$ then $\Pr(\max(X, Y) - \min(X, Y) > t) = 1$. If $t > 0$ then

$$\begin{aligned}\Pr(\max(X, Y) - \min(X, Y) > t) =&\Pr(|X - Y| > t)\\ =&1 - \Phi\left(\frac{t-\mu}{\sqrt{2}}\right) + \Phi\left(\frac{-t-\mu}{\sqrt{2}}\right)\end{aligned}$$

Note that $X - Y$ is normal with mean μ and variance 2 ■

Example 40.8 *(Order statistics)*
Suppose $X_1, \cdots, X_n$ are independent and identically distributed random variables with cdf $F_X(x)$. Define U and L as

$$\begin{aligned}U =&\max\{X_1, X_2, \cdots, X_n\}\\ L =&\min\{X_1, X_2, \cdots, X_n\}\end{aligned}$$

(a) Find the cdf of U.
(b) Find the cdf of L.
(c) Are U and L independent?

Solution.
(a) First note the following equivalence of events

$$\{U \leq u\} \Leftrightarrow \{X_1 \leq u, X_2 \leq u, \cdots, X_n \leq u\}.$$

Thus,

$$\begin{aligned}F_U(u) =&\Pr(U \leq u) = \Pr(X_1 \leq u, X_2 \leq u, \cdots, X_n \leq u)\\ =&\Pr(X_1 \leq u)\Pr(X_2 \leq u) \cdots \Pr(X_n \leq u) = (F_X(x))^n\end{aligned}$$

(b) Note the following equivalence of events

$$\{L > l\} \Leftrightarrow \{X_1 > l, X_2 > l, \cdots, X_n > l\}.$$

Thus,

$$\begin{aligned}F_L(l) =&\Pr(L \leq l) = 1 - \Pr(L > l)\\ =&1 - \Pr(X_1 > l, X_2 > l, \cdots, X_n > l)\\ =&1 - \Pr(X_1 > l)\Pr(X_2 > l) \cdots \Pr(X_n > l)\\ =&1 - (1 - F_X(x))^n\end{aligned}$$

(c) No. First note that $\Pr(L > l) = 1 - F_L(l)$. From the definition of cdf there must be a number l_0 such that $F_L(l_0) \neq 1$. Thus, $\Pr(L > l_0) \neq 0$. But $\Pr(L > l_0|U \leq u) = 0$ for any $u < l_0$. This shows that $\Pr(L > l_0|U \leq u) \neq \Pr(L > l_0)$ ■

Remark 40.1
L defined in the previous example is referred to as the **first order statistics**. U is referred to as the n^{th} **order statistics**.

Practice Problems

Problem 40.1
Let X and Y be random variables with joint pdf given by

$$f_{XY}(x,y) = \begin{cases} e^{-(x+y)} & 0 \le x, y \\ 0 & \text{otherwise.} \end{cases}$$

(a) Are X and Y independent?
(b) Find $\Pr(X < Y)$.
(c) Find $\Pr(X < a)$.

Problem 40.2
The random vector (X, Y) is said to be **uniformly distributed** over a region R in the plane if, for some constant c, its joint pdf is

$$f_{XY}(x,y) = \begin{cases} c & (x,y) \in R \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that $c = \frac{1}{A(R)}$ where $A(R)$ is the area of the region R.
(b) Suppose that $R = \{(x,y) : -1 \le x \le 1, -1 \le y \le 1\}$. Show that X and Y are independent, with each being distributed uniformly over $(-1, 1)$.
(c) Find $\Pr(X^2 + Y^2 \le 1)$.

Problem 40.3
Let X and Y be random variables with joint pdf given by

$$f_{XY}(x,y) = \begin{cases} 6(1-y) & 0 \le x \le y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $\Pr(X \le \frac{3}{4}, Y \ge \frac{1}{2})$.
(b) Find $f_X(x)$ and $f_Y(y)$.
(c) Are X and Y independent?

Problem 40.4
Let X and Y have the joint pdf given by

$$f_{XY}(x,y) = \begin{cases} kxy & 0 \le x, y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find k.
(b) Find $f_X(x)$ and $f_Y(y)$.
(c) Are X and Y independent?

Problem 40.5
Let X and Y have joint density

$$f_{XY}(x,y) = \begin{cases} kxy^2 & 0 \le x, y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find k.
(b) Compute the marginal densities of X and of Y.
(c) Compute $\Pr(Y > 2X)$.
(d) Compute $\Pr(|X - Y| < 0.5)$.
(e) Are X and Y independent?

Problem 40.6
Suppose the joint density of random variables X and Y is given by

$$f_{XY}(x,y) = \begin{cases} kx^2y^{-3} & 1 \le x, y \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find k.
(b) Are X and Y independent?
(c) Find $\Pr(X > Y)$.

Problem 40.7
Let X and Y be continuous random variables, with the joint probability density function

$$f_{XY}(x,y) = \begin{cases} \frac{3x^2+2y}{24} & 0 \le x, y \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $f_X(x)$ and $f_Y(y)$.
(b) Are X and Y independent?
(c) Find $\Pr(X + 2Y < 3)$.

Problem 40.8
Let X and Y have joint density

$$f_{XY}(x,y) = \begin{cases} \frac{4}{9} & x \le y \le 3 - x, 0 \le x \\ 0 & \text{otherwise.} \end{cases}$$

(a) Compute the marginal densities of X and Y.
(b) Compute $\Pr(Y > 2X)$.
(c) Are X and Y independent?

Problem 40.9 ‡
A study is being conducted in which the health of two independent groups of ten policyholders is being monitored over a one-year period of time. Individual participants in the study drop out before the end of the study with probability 0.2 (independently of the other participants).
What is the probability that at least 9 participants complete the study in one of the two groups, but not in both groups?

Problem 40.10 ‡
The waiting time for the first claim from a good driver and the waiting time for the first claim from a bad driver are independent and follow exponential distributions with means 6 years and 3 years, respectively.
What is the probability that the first claim from a good driver will be filed within 3 years and the first claim from a bad driver will be filed within 2 years?

Problem 40.11 ‡
An insurance company sells two types of auto insurance policies: Basic and Deluxe. The time until the next Basic Policy claim is an exponential random variable with mean two days. The time until the next Deluxe Policy claim is an independent exponential random variable with mean three days.
What is the probability that the next claim will be a Deluxe Policy claim?

Problem 40.12 ‡
Two insurers provide bids on an insurance policy to a large company. The bids must be between 2000 and 2200 . The company decides to accept the lower bid if the two bids differ by 20 or more. Otherwise, the company will consider the two bids further. Assume that the two bids are independent and are both uniformly distributed on the interval from 2000 to 2200.
Determine the probability that the company considers the two bids further.

Problem 40.13 ‡
A family buys two policies from the same insurance company. Losses under the two policies are independent and have continuous uniform distributions on the interval from 0 to 10. One policy has a deductible of 1 and the other has a deductible of 2. The family experiences exactly one loss under each policy.
Calculate the probability that the total benefit paid to the family does not exceed 5.

Problem 40.14 ‡
In a small metropolitan area, annual losses due to storm, fire, and theft are assumed to be independent, exponentially distributed random variables with respective means 1.0, 1.5, and 2.4 .
Determine the probability that the maximum of these losses exceeds 3.

Problem 40.15 ‡
A device containing two key components fails when, and only when, both components fail. The lifetimes, X and Y, of these components are independent with common density function $f(t) = e^{-t}, t > 0$. The cost, Z, of operating the device until failure is $2X + Y$.
Find the probability density function of Z.

Problem 40.16 ‡
A company offers earthquake insurance. Annual premiums are modeled by an exponential random variable with mean 2. Annual claims are modeled by an exponential random variable with mean 1. Premiums and claims are independent. Let X denote the ratio of claims to premiums.
What is the density function of X?

Problem 40.17
Let X and Y be independent continuous random variables with common density function

$$f_X(x) = f_Y(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

What is $\Pr(X^2 \geq Y^3)$?

Problem 40.18
Suppose that discrete random variables X and Y each take only the values 0 and 1. It is known that $\Pr(X = 0|Y = 1) = 0.6$ and $\Pr(X = 1|Y = 0) = 0.7$. Is it possible that X and Y are independent? Justify your conclusion.

Problem 40.19
Let X and Y be two discrete random variables with joint distribution given by the following table.

Y\ X	1	5
2	$\theta_1 + \theta_2$	$\theta_1 + 2\theta_2$
4	$\theta_1 + 2\theta_2$	$\theta_1 + \theta_2$

We assume that $-0.25 \leq \theta_1 \leq 2.5$ and $0 \leq \theta_2 \leq 0.35$. Find θ_1 and θ_2 when X and Y are independent.

41 Sum of Two Independent Random Variables: Discrete Case

In this section we turn to the important question of determining the distribution of a sum of independent random variables in terms of the distributions of the individual constituents. In this section, we consider only sums of discrete random variables, reserving the case of continuous random variables for the next section. We consider here only discrete random variables whose values are nonnegative integers. Their distribution mass functions are then defined on these integers.

Suppose X and Y are two independent discrete random variables with pmf $p_X(x)$ and $p_Y(y)$ respectively. We would like to determine the pmf of the random variable $X+Y$. To do this, we note first that for any nonnegative integer n we have

$$\{X+Y=n\}=\bigcup_{k=0}^{n} A_k$$

where $A_k=\{X=k\}\cap\{Y=n-k\}$. Note that $A_i\cap A_j=\emptyset$ for $i\neq j$. Since the A_i's are pairwise disjoint and X and Y are independent, we have

$$\Pr(X+Y=n)=\sum_{k=0}^{n}\Pr(X=k)\Pr(Y=n-k).$$

Thus,

$$p_{X+Y}(n)=p_X(n)*p_Y(n)$$

where $p_X(n)*p_Y(n)$ is called the **convolution** of p_X and p_Y.

Example 41.1

A die is rolled twice. Let X and Y be the outcomes, and let $Z=X+Y$ be the sum of these outcomes. Find the probability mass function of Z.

Solution. Note that X and Y have the common pmf :

x	1	2	3	4	5	6
p_X	1/6	1/6	1/6	1/6	1/6	1/6

The probability mass function of Z is then the convolution of p_X with itself. Thus,

$$
\begin{aligned}
\Pr(Z=2) =& p_X(1)p_X(1) = \frac{1}{36} \\
\Pr(Z=3) =& p_X(1)p_X(2) + p_X(2)p_X(1) = \frac{2}{36} \\
\Pr(Z=4) =& p_X(1)p_X(3) + p_X(2)p_X(2) + p_X(3)p_X(1) = \frac{3}{36}
\end{aligned}
$$

Continuing in this way we would find $\Pr(Z=5) = 4/36, \Pr(Z=6) = 5/36, \Pr(Z=7) = 6/36, \Pr(Z=8) = 5/36, \Pr(Z=9) = 4/36, \Pr(Z=10) = 3/36, \Pr(Z=11) = 2/36$, and $\Pr(Z=12) = 1/36$ ■

Example 41.2

Let X and Y be two independent Poisson random variables with respective parameters λ_1 and λ_2. Compute the pmf of $X+Y$.

Solution.

For every positive integer we have

$$\{X+Y=n\} = \bigcup_{k=0}^{n} A_k$$

where $A_k = \{X=k, Y=n-k\}$ for $0 \le k \le n$. Moreover, $A_i \cap A_j = \emptyset$ for $i \ne j$. Thus,

$$
\begin{aligned}
p_{X+Y}(n) =& \Pr(X+Y=n) = \sum_{k=0}^{n} \Pr(X=k, Y=n-k) \\
=& \sum_{k=0}^{n} \Pr(X=k)\Pr(Y=n-k) \\
=& \sum_{k=0}^{n} e^{-\lambda_1} \frac{\lambda_1^k}{k!} e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} \\
=& e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{n} \frac{\lambda_1^k \lambda_2^{n-k}}{k!(n-k)!} \\
=& \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \lambda_1^k \lambda_2^{n-k} \\
=& \frac{e^{-(\lambda_1+\lambda_2)}}{n!} (\lambda_1+\lambda_2)^n
\end{aligned}
$$

Thus, $X+Y$ is a Poisson random variable with parameter $\lambda_1+\lambda_2$ ■

Example 41.3
Let X and Y be two independent binomial random variables with respective parameters (n,p) and (m,p). Compute the pmf of $X+Y$.

Solution.
X represents the number of successes in n independent trials, each of which results in a success with probability p; similarly, Y represents the number of successes in m independent trials, each of which results in a success with probability p. Hence, as X and Y are assumed to be independent, it follows that $X+Y$ represents the number of successes in $n+m$ independent trials, each of which results in a success with probability p. So $X+Y$ is a binomial random variable with parameters $(n+m,p)$ ■

Example 41.4
Two bias coins are being flipped repeatedly. The probability that coin 1 comes up heads is $\frac{1}{4}$, while that of coin 2 is $\frac{3}{4}$. Each coin is being flipped until a head comes up. What is the pmf of the total number of flips until both coins come up heads?

Solution.
Let X and Y be the number of flips of coins 1 and 2 to come up heads for the first time. Then, $X+Y$ is the total number of flips until both coins come up heads for the first time. The random variables X and Y are independent geometric random variables with parameters $1/4$ and $3/4$, respectively. By convolution, we have

$$\begin{aligned}p_{X+Y}(n)=&\sum_{k=1}^{n-1}\frac{1}{4}\left(\frac{3}{4}\right)^{k-1}\frac{3}{4}\left(\frac{1}{4}\right)^{n-k-1}\\=&\frac{1}{4^n}\sum_{k=1}^{n-1}3^k=\frac{3}{2}\frac{3^{n-1}-1}{4^n}\ ■\end{aligned}$$

Practice Problems

Problem 41.1
Let X and Y be two independent discrete random variables with probability mass functions defined in the tables below. Find the probability mass function of $Z = X + Y$.

x	0	1	2	3
$p_X(x)$	0.10	0.20	0.30	0.40

y	0	1	2
$p_Y(y)$	0.25	0.40	0.35

Problem 41.2
Suppose X and Y are two independent binomial random variables with respective parameters $(20, 0.2)$ and $(10, 0.2)$. Find the pmf of $X + Y$.

Problem 41.3
Let X and Y be independent random variables each geometrically distributed with parameter p, i.e.

$$p_X(n) = p_Y(n) = \begin{cases} p(1-p)^{n-1} & n = 1, 2, \cdots \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability mass function of $X + Y$.

Problem 41.4
Consider the following two experiments: the first has outcome X taking on the values 0, 1, and 2 with equal probabilities; the second results in an (independent) outcome Y taking on the value 3 with probability 1/4 and 4 with probability 3/4. Find the probability mass function of $X + Y$.

Problem 41.5 ‡
An insurance company determines that N, the number of claims received in a week, is a random variable with $P[N = n] = \frac{1}{2^{n+1}}$, where $n \geq 0$. The company also determines that the number of claims received in a given week is independent of the number of claims received in any other week.
Determine the probability that exactly seven claims will be received during a given two-week period.

Problem 41.6
Suppose X and Y are independent, each having Poisson distribution with means 2 and 3, respectively. Let $Z = X + Y$. Find $\Pr(X + Y = 1)$.

Problem 41.7
Suppose that X has Poisson distribution with parameter λ and that Y has geometric distribution with parameter p and is independent of X. Find simple formulas in terms of λ and p for the following probabilities. (The formulas should not involve an infinite sum.)
(a) $\Pr(X + Y = 2)$
(b) $\Pr(Y > X)$

Problem 41.8
Let X and Y be two independent random variables with common pmf given by

x	0	1	2	y
$p_X(x)$	0.5	0.25	0.25	$p_Y(y)$

Find the probability mass function of $X + Y$.

Problem 41.9
Let X and Y be two independent random variables with pmfs given by

$$p_X(x) = \begin{cases} \frac{1}{3} & x = 1, 2, 3 \\ 0 & \text{otherwise} \end{cases}$$

$$p_Y(y) = \begin{cases} \frac{1}{2} & y = 0 \\ \frac{1}{3} & y = 1 \\ \frac{1}{6} & y = 2 \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability mass function of $X + Y$.

Problem 41.10
Let X and Y be two independent identically distributed geometric distributions with parameter p. Show that $X + Y$ is a negative binomial distribution with parameters $(2, p)$.

Problem 41.11
Let X, Y, Z be independent Poisson random variables with $E(X) = 3, E(Y) = 1$, and $E(Z) = 4$. What is $\Pr(X + Y + Z \leq 1)$?

Problem 41.12
The number of phone calls received by an operator in 5-minute period follows a Poisson distribution with a mean of λ. Find the probability that the total number of phone calls received in 10 randomly selected 5-minute periods is 10.

42 Sum of Two Independent Random Variables: Contniuous Case

In this section, we consider the continuous version of the problem posed in Section 41: How are sums of independent continuous random variables distributed?

Example 42.1
Let X and Y be two random variables with joint probability density

$$f_{XY}(x,y) = \begin{cases} 6e^{-3x-2y} & x > 0, y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Find the probability density of $Z = X + Y$.

Solution.
Integrating the joint probability density over the shaded region of Figure 42.1, we get

$$F_Z(a) = \Pr(Z \le a) = \int_0^a \int_0^{a-y} 6e^{-3x-2y} dxdy = 1 + 2e^{-3a} - 3e^{-2a}$$

and differentiating with respect to a we find

$$f_Z(a) = 6(e^{-2a} - e^{-3a})$$

for $a > 0$ and 0 elsewhere ■

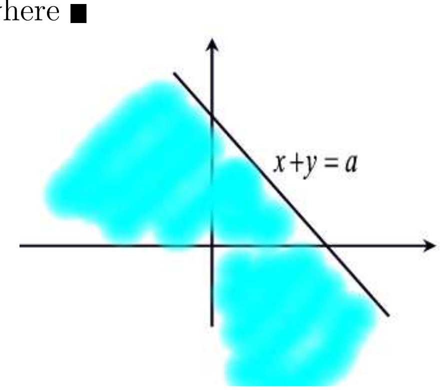


Figure 42.1

The above process can be generalized with the use of convolutions which we define next. Let X and Y be two continuous random variables with probability density functions $f_X(x)$ and $f_Y(y)$, respectively. Assume that both $f_X(x)$ and $f_Y(y)$ are defined for all real numbers. Then the **convolution** $f_X * f_Y$ of f_X and f_Y is the function given by

$$\begin{aligned}(f_X * f_Y)(a) &= \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy \\ &= \int_{-\infty}^{\infty} f_Y(a-x)f_X(x)dx\end{aligned}$$

This definition is analogous to the definition, given for the discrete case, of the convolution of two probability mass functions. Thus it should not be surprising that if X and Y are independent, then the probability density function of their sum is the convolution of their densities.

Theorem 42.1
Let X and Y be two independent random variables with density functions $f_X(x)$ and $f_Y(y)$ defined for all x and y. Then the sum $X+Y$ is a random variable with density function $f_{X+Y}(a)$, where f_{X+Y} is the convolution of f_X and f_Y.

Proof.
The cumulative distribution function is obtained as follows:

$$\begin{aligned}F_{X+Y}(a) =&\Pr(X+Y \le a) = \iint_{x+y\le a} f_X(x)f_Y(y)dxdy \\ =& \int_{-\infty}^{\infty}\int_{-\infty}^{a-y} f_X(x)f_Y(y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{a-y} f_X(x)dxf_Y(y)dy \\ =& \int_{-\infty}^{\infty} F_X(a-y)f_Y(y)dy\end{aligned}$$

Differentiating the previous equation with respect to a we find

$$\begin{aligned}f_{X+Y}(a) =& \frac{d}{da}\int_{-\infty}^{\infty} F_X(a-y)f_Y(y)dy \\ =& \int_{-\infty}^{\infty} \frac{d}{da}F_X(a-y)f_Y(y)dy \\ =& \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy \\ =& (f_X * f_Y)(a) \blacksquare\end{aligned}$$

Example 42.2
Let X and Y be two independent random variables uniformly distributed on $[0, 1]$. Compute the probability density function of $X + Y$.

Solution.
Since

$$f_X(a) = f_Y(a) = \begin{cases} 1 & 0 \le a \le 1 \\ 0 & \text{otherwise} \end{cases}$$

by the previous theorem

$$f_{X+Y}(a) = \int_0^1 f_X(a - y)dy.$$

Now the integrand is 0 unless $0 \le a - y \le 1$(i.e. unless $a - 1 \le y \le a$) and then it is 1. So if $0 \le a \le 1$ then

$$f_{X+Y}(a) = \int_0^a dy = a.$$

If $1 < a < 2$ then

$$f_{X+Y}(a) = \int_{a-1}^1 dy = 2 - a.$$

Hence,

$$f_{X+Y}(a) = \begin{cases} a & 0 \le a \le 1 \\ 2 - a & 1 < a < 2 \\ 0 & \text{otherwise} \ \blacksquare \end{cases}$$

Example 42.3
Let X and Y be two independent exponential random variables with common parameter λ. Compute $f_{X+Y}(a)$.

Solution.
We have

$$f_X(a) = f_Y(a) = \begin{cases} \lambda e^{-\lambda a} & 0 \le a \\ 0 & \text{otherwise.} \end{cases}$$

If $a \ge 0$ then

$$\begin{aligned} f_{X+Y}(a) =& \int_{-\infty}^{\infty} f_X(a - y)f_Y(y)dy \\ =& \lambda^2 \int_0^a e^{-\lambda a}dy = a\lambda^2 e^{-\lambda a}. \end{aligned}$$

If $a < 0$ then $f_{X+Y}(a) = 0$. Hence,

$$f_{X+Y}(a) = \begin{cases} a\lambda^2 e^{-\lambda a} & 0 \le a \\ 0 & \text{otherwise} \end{cases} \blacksquare$$

Example 42.4
Let X and Y be two independent random variables, each with the standard normal density. Compute $f_{X+Y}(a)$.

Solution.
We have

$$f_X(a) = f_Y(a) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

By Theorem 42.1 we have

$$\begin{aligned} f_{X+Y}(a) =& \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{(a-y)^2}{2}} e^{-\frac{y^2}{2}} dy \\ =& \frac{1}{2\pi} e^{-\frac{a^2}{4}} \int_{-\infty}^{\infty} e^{-(y-\frac{a}{2})^2} dy \\ =& \frac{1}{2\pi} e^{-\frac{a^2}{4}} \sqrt{\pi} \left[\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-w^2} dw \right], \quad w = y - \frac{a}{2}. \end{aligned}$$

The expression in the brackets equals 1, since it is the integral of the normal density function with $\mu = 0$ and $\sigma = \frac{1}{\sqrt{2}}$. Hence,

$$f_{X+Y}(a) = \frac{1}{\sqrt{4\pi}} e^{-\frac{a^2}{4}} \blacksquare$$

Example 42.5
Let X and Y be two independent gamma random variables with respective parameters (s, λ) and (t, λ). Show that $X + Y$ is a gamma random variable with parameters $(s + t, \lambda)$.

Solution.
We have

$$f_X(a) = \frac{\lambda e^{-\lambda a} (\lambda a)^{s-1}}{\Gamma(s)} \quad \text{and} \quad f_Y(a) = \frac{\lambda e^{-\lambda a} (\lambda a)^{t-1}}{\Gamma(t)}$$

By Theorem 42.1 we have

$$\begin{aligned}
f_{X+Y}(a) =&\frac{1}{\Gamma(s)\Gamma(t)}\int_0^a \lambda e^{-\lambda(a-y)}[\lambda(a-y)]^{s-1}\lambda e^{-\lambda y}(\lambda y)^{t-1}dy\\
=&\frac{\lambda^{s+t}e^{-\lambda a}}{\Gamma(s)\Gamma(t)}\int_0^a (a-y)^{s-1}y^{t-1}dy\\
=&\frac{\lambda^{s+t}e^{-\lambda a}a^{s+t-1}}{\Gamma(s)\Gamma(t)}\int_0^1 (1-x)^{s-1}x^{t-1}dx, \quad x=\frac{y}{a}.
\end{aligned}$$

Using the fact that

$$\int_0^1 (1-x)^{s-1}x^{t-1}dx = \frac{\Gamma(s)\Gamma(t)}{\Gamma(s+t)}.$$

we can write

$$f_{X+Y}(a) = \frac{\lambda e^{-\lambda a}(\lambda a)^{s+t-1}}{\Gamma(s+t)} \blacksquare$$

Example 42.6
The joint distribution function of X and Y is given by

$$f_{XY}(x,y) = \begin{cases} \frac{3}{11}(5x+y) & x,y>0, \quad x+2y<2 \\ 0 & \text{elsewhere} \end{cases}$$

Find the probability density of $Z = X+Y$.

Solution.
Note first that the region of integration is the interior of the triangle with vertices at $(0,0), (0,1)$, and $(2,0)$. From Figure 42.2, we see that $F(a)=0$ if $a<0$. If $0\le a<1$ then

$$F_Z(a) = \Pr(Z\le a) = \int_0^a \int_0^{a-y} \frac{3}{11}(5x+y)dxdy = \frac{3}{11}a^3.$$

If $1\le a<2$ then the two lines $x+y=a$ and $x+2y=2$ intersect at $(2a-2, 2-a)$. In this case,

$$\begin{aligned}
F_Z(a) =&\Pr(Z\le a)\\
=&\int_0^{2-a}\int_0^{a-y}\frac{3}{11}(5x+y)dxdy + \int_{2-a}^1\int_0^{2-2y}\frac{3}{11}(5x+y)dxdy\\
=&\frac{3}{11}\left(-\frac{7}{3}a^3+9a^2-8a+\frac{7}{3}\right).
\end{aligned}$$

If $a \geq 2$ then $F_Z(a)$ is the area of the shaded triangle which is equal to 1. Differentiating with respect to a we find

$$f_Z(a) = \begin{cases} \frac{9}{11}a^2 & 0 < a \leq 1 \\ \frac{3}{11}(-7a^2 + 18a - 8) & 1 < a < 2 \\ 0 & \text{elsewhere} \end{cases} \blacksquare$$

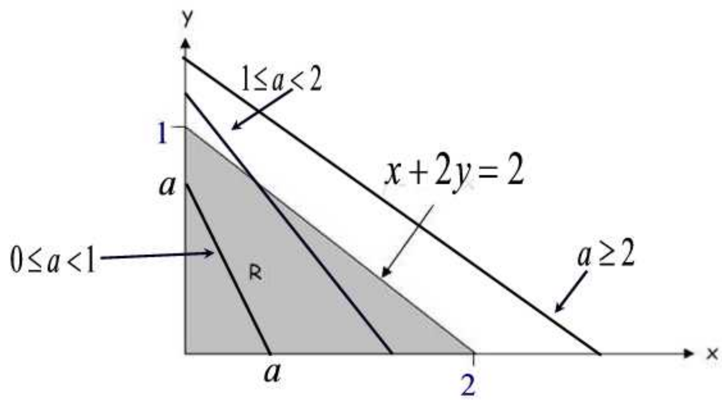


Figure 42.2

Practice Problems

Problem 42.1
Let X be an exponential random variable with parameter λ and Y be an exponential random variable with parameter 2λ independent of X. Find the probability density function of $X+Y$.

Problem 42.2
Let X be an exponential random variable with parameter λ and Y be a uniform random variable on [0,1] independent of X. Find the probability density function of $X+Y$.

Problem 42.3
Let X and Y be two independent random variables with probability density functions (p.d.f.) , f_X and f_Y respectively. Find the pdf of $X+2Y$.

Problem 42.4
Consider two independent random variables X and Y. Let $f_X(x) = 1 - \frac{x}{2}$ if $0 \leq x \leq 2$ and 0 otherwise. Let $f_Y(y) = 2 - 2y$ for $0 \leq y \leq 1$ and 0 otherwise. Find the probability density function of $X+Y$.

Problem 42.5
Let X and Y be two independent and identically distributed random variables with common density function

$$f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the probability density function of $X+Y$.

Problem 42.6
Let X and Y be independent exponential random variables with pairwise distinct respective parameters α and β. Find the probability density function of $X+Y$.

Problem 42.7
Let X and Y be two random variables with common pdf

$$f_X(t) = f_Y(t) = \begin{cases} e^{-t} & t > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find the density function of $W = 2X + Y$.

Problem 42.8
Let X and Y be independent random variables with density functions

$$f_X(x) = \begin{cases} \frac{1}{2} & -1 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{2} & 3 \le y \le 5 \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability density function of $X + Y$.

Problem 42.9
Let X and Y be independent random variables with density functions

$$f_X(x) = \begin{cases} \frac{1}{2} & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{y}{2} & 0 < y < 2 \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability density function of $X + Y$.

Problem 42.10
Let X and Y be independent random variables with density functions

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

and

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}.$$

Find the probability density function of $X + Y$.

Problem 42.11
Let X have a uniform distribution on the interval $(1, 3)$. What is the probability that the sum of 2 independent observations of X is greater than 5?

Problem 42.12
Let X and Y be two independent exponential random variables each with mean 1. Find the pdf of $Z = X + Y$.

Problem 42.13
X_1 and X_2 are independent exponential random variables each with a mean of 1. Find $\Pr(X_1 + X_2 < 1)$.

43 Conditional Distributions: Discrete Case

Recall that for any two events E and F the conditional probability of E given F is defined by

$$\Pr(E|F) = \frac{\Pr(E \cap F)}{\Pr(F)}$$

provided that $\Pr(F) > 0$.
In a similar way, if X and Y are discrete random variables then we define the **conditional probability mass function** of X given that $Y = y$ by

$$p_{X|Y}(x|y) = \Pr(X = x|Y = y) = \frac{\Pr(X = x, Y = y)}{\Pr(Y = y)} = \frac{p_{XY}(x, y)}{p_Y(y)} \quad (43.1)$$

provided that $p_Y(y) > 0$.

Example 43.1
Two coins are being tossed repeatedly. The tossing of each coin stops when the coin comes up a head.
(a) Find the probability that the two coins come up heads at the same time.
(b) Find the conditional distribution of the number of coin tosses given that the two coins come up heads simultaneously.

Solution.
(a) Let X be the number of tosses of the first coin before getting a head, and Y be the number of tosses of the second coin before getting a head. So X and Y are independent identically distributed geometric random variables with parameter $p = \frac{1}{2}$. Thus,

$$\begin{aligned}\Pr(X = Y) =& \sum_{k=1}^{\infty} \Pr(X = k, Y = k) = \sum_{k=1}^{\infty} \Pr(X = k)\Pr(Y = k)\\ =& \sum_{k=1}^{\infty} \frac{1}{4^k} = \frac{1}{3}\end{aligned}$$

(b) Notice that given the event $[X = Y]$ the number of coin tosses is well defined and it is X (or Y). So for any $k \geq 1$ we have

$$\Pr(X = k|Y = k) = \frac{\Pr(X = k, Y = k)}{\Pr(X = Y)} = \frac{\frac{1}{4^k}}{\frac{1}{3}} = \frac{3}{4}\left(\frac{1}{4}\right)^{k-1}.$$

Thus given $[X = Y]$, the number of tosses follows a geometric distribution with parameter $p = \frac{3}{4}$ ■

Sometimes it is not the joint distribution that is known, but rather, for each y, one knows the conditional distribution of X given $Y = y$. If one also knows the distribution of Y, then one can recover the joint distribution using (43.1). We also mention one more use of (43.1):

$$p_X(x) = \sum_y p_{XY}(x, y) = \sum_y p_{X|Y}(x|y)p_Y(y). \tag{43.2}$$

Thus, given the conditional distribution of X given $Y = y$ for each possible value y, and the (marginal) distribution of Y, one can compute the (marginal) distribution of X, using (43.2).
The **conditional cumulative distribution** of X given that $Y = y$ is defined by

$$F_{X|Y}(x|y) = \Pr(X \le x|Y = y) = \sum_{a \le x} p_{X|Y}(a|y).$$

Note that if X and Y are independent, then the conditional mass function and the conditional distribution function are the same as the unconditional ones. This follows from the next theorem.

Theorem 43.1
If X and Y are independent and $p_Y(y) > 0$ then

$$p_{X|Y}(x|y) = p_X(x).$$

Proof.
We have

$$\begin{aligned} p_{X|Y}(x|y) =& \Pr(X = x|Y = y) \\ =& \frac{\Pr(X = x, Y = y)}{\Pr(Y = y)} \\ =& \frac{\Pr(X = x)\Pr(Y = y)}{\Pr(Y = y)} \\ =& \Pr(X = x) = p_X(x) \ \blacksquare \end{aligned}$$

Example 43.2
Given the following table.

X\ Y	Y=1	Y=2	Y=3	$p_X(x)$
X=1	.01	.20	.09	.3
X=2	.07	.00	.03	.1
X=3	.09	.05	.06	.2
X=4	.03	.25	.12	.4
$p_Y(y)$	.2	.5	.3	1

Find $p_{X|Y}(x|y)$ where $Y = 2$.

Solution.

$$\begin{aligned}
p_{X|Y}(1|2) =& \frac{p_{XY}(1,2)}{p_Y(2)} = \frac{.2}{.5} = 0.4\\
p_{X|Y}(2|2) =& \frac{p_{XY}(2,2)}{p_Y(2)} = \frac{0}{.5} = 0\\
p_{X|Y}(3|2) =& \frac{p_{XY}(3,2)}{p_Y(2)} = \frac{.05}{.5} = 0.1\\
p_{X|Y}(4|2) =& \frac{p_{XY}(4,2)}{p_Y(2)} = \frac{.25}{.5} = 0.5\\
p_{X|Y}(x|2) =& \frac{p_{XY}(x,2)}{p_Y(2)} = \frac{0}{.5} = 0, \quad x > 4 \blacksquare
\end{aligned}$$

Example 43.3
If X and Y are independent Poisson random variables with respective parameters λ_1 and λ_2, calculate the conditional distribution of X, given that $X + Y = n$.

Solution.
We have

$$\begin{aligned}
\Pr(X=k|X+Y=n)=&\frac{\Pr(X=k,X+Y=n)}{\Pr(X+Y=n)}\\
=&\frac{\Pr(X=k,Y=n-k)}{\Pr(X+Y=n)}\\
=&\frac{\Pr(X=k)\Pr(Y=n-k))}{\Pr(X+Y=n)}\\
=&\frac{e^{-\lambda_1}\lambda_1^k}{k!}\frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!}\left[\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!}\right]^{-1}\\
=&\frac{n!}{k!(n-k)!}\frac{\lambda_1^k\lambda_2^{n-k}}{(\lambda_1+\lambda_2)^n}\\
=&\left(\begin{array}{c}n\\k\end{array}\right)\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}
\end{aligned}$$

In other words, the conditional mass distribution function of X given that $X+Y=n$, is the binomial distribution with parameters n and $\frac{\lambda_1}{\lambda_1+\lambda_2}$ ■

Practice Problems

Problem 43.1
Given the following table.

X\ Y	Y=0	Y=1	$p_X(x)$
X=0	.4	.1	.5
X=1	.2	.3	.5
$p_Y(y)$	.6	.4	1

Find $p_{X|Y}(x|y)$ where $Y = 1$.

Problem 43.2
Let X be a random variable with range the set $\{1, 2, 3, 4, 5\}$ and Y be a random variable with range the set $\{1, 2, \cdots, X\}$.
(a) Find $p_{XY}(x, y)$.
(b) Find $p_{X|Y}(x|y)$.
(c) Are X and Y independent?

Problem 43.3
The following is the joint distribution function of X and Y.

$X\backslash Y$	4	3	2	$p_X(x)$
5	0.1	0.05	0	0.15
4	0.15	0.15	0	0.3
3	0.10	0.15	0.10	0.35
2	0	0.05	0.10	0.15
1	0	0	0.05	0.05
$p_Y(y)$	0.35	0.40	0.25	1

Find $\Pr(X|Y = 4)$ for $X = 3, 4, 5$.

Problem 43.4
A fair coin is tossed 4 times. Let the random variable X denote the number of heads in the first 3 tosses, and let the random variable Y denote the number of heads in the last 3 tosses. The joint pmf is given by the following table

$X\backslash Y$	0	1	2	3	$p_X(x)$
0	1/16	1/16	0	0	2/16
1	1/16	3/16	2/16	0	6/16
2	0	2/16	3/16	1/16	6/16
3	0	0	1/16	1/16	2/16
$p_Y(y)$	2/16	6/16	6/16	2/16	1

What is the conditional pmf of the number of heads in the first 3 coin tosses given exactly 1 head was observed in the last 3 tosses?

Problem 43.5
Two dice are rolled. Let X and Y denote, respectively, the largest and smallest values obtained. Compute the conditional mass function of Y given $X = x$, for $x = 1, 2, \cdots, 6$. Are X and Y independent?

Problem 43.6
Let X and Y be discrete random variables with joint probability function

$$p_{XY}(x,y) = \begin{cases} \frac{n!y^x(pe^{-1})^y(1-p)^{n-y}}{y!(n-y)!x!} & y = 0, 1, \cdots, n\ ; x = 0, 1, \cdots \\ 0 & \text{otherwise} \end{cases}$$

(a) Find $p_Y(y)$.
(b) Find the conditional probability distribution of X, given $Y = y$. Are X and Y independent? Justify your answer.

Problem 43.7
Let X and Y have the joint probability function $p_{XY}(x,y)$ described as follows:

X\ Y	0	1	$p_X(x)$
0	1/18	3/18	4/18
1	4/18	3/18	7/18
2	6/18	1/18	7/18
$p_Y(y)$	11/18	7/18	1

Find $p_{X|Y}(x|y)$ and $p_{Y|X}(y|x)$.

Problem 43.8
Let X and Y be random variables with joint probability mass function

$$p_{XY}(x,y) = c(1 - 2^{-x})^y$$

where $x = 0, 1, \cdots, N-1$ and $y = 0, 1, 2, \cdots$
(a) Find c.
(b) Find $p_X(x)$.
(c) Find $p_{Y|X}(y|x)$, the conditional probability mass function of Y given $X = x$.

Problem 43.9
Let X and Y be identically independent Poisson random variables with paramter λ. Find $\Pr(X = k|X + Y = n)$.

Problem 43.10
If two cards are randomly drawn (without replacement) from an ordinary deck of 52 playing cards, Y is the number of aces obtained in the first draw and X is the total number of aces obtained in both draws, find
(a) the joint probability distribution of X and Y;
(b) the marginal distribution of Y;
(c) the conditional distribution of X given $Y = 1$.

Problem 43.11 ‡
Let N_1 and N_2 represent the numbers of claims submitted to a life insurance company in April and May, respectively. The joint probability function of N_1 and N_2 is

$$\Pr(n_1, n_2) = \begin{cases} \frac{3}{4}\left(\frac{1}{4}\right)^{n_1-1} e^{-n_1}(1-e^{-n_1})^{n_2-1}, & \text{for } n_1 = 1, 2, 3, \cdots \text{ and } n_2 = 1, 2, 3, \cdots \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the expected number of claims that will be submitted to the company in May if exactly 2 claims were submitted in April.

44 Conditional Distributions: Continuous Case

In this section, we develop the distribution of X given Y when both are continuous random variables. Unlike the discrete case, we cannot use simple conditional probability to define the conditional probability of an event given $Y=y$, because the conditioning event has probability 0 for any y. However, we motivate our approach by the following argument.
Suppose X and Y are two continuous random variables with joint density $f_{XY}(x,y)$. Let $f_{X|Y}(x|y)$ denote the probability density function of X given that $Y=y$. We define

$$\Pr(a<X<b|Y=y)=\int_a^b f_{X|Y}(x|y)dx.$$

Then for δ very small we have (See Remark 30.1)

$$\Pr(x\le X\le x+\delta|Y=y)\approx \delta f_{X|Y}(x|y).$$

On the other hand, for small ϵ we have

$$\begin{aligned}\Pr(x\le X\le x+\delta|Y=y)\approx&\Pr(x\le X\le x+\delta|y\le Y\le y+\epsilon)\\=&\frac{\Pr(x\le X\le x+\delta, y\le Y\le y+\epsilon)}{\Pr(y\le Y\le y+\epsilon)}\\\approx&\frac{\delta\epsilon f_{XY}(x,y)}{\epsilon f_Y(y)}.\end{aligned}$$

In the limit, as ϵ tends to 0, we are left with

$$\delta f_{X|Y}(x|y)\approx\frac{\delta f_{XY}(x,y)}{f_Y(y)}.$$

This suggests the following definition. The **conditional density function** of X given $Y=y$ is

$$f_{X|Y}(x|y)=\frac{f_{XY}(x,y)}{f_Y(y)}$$

provided that $f_Y(y)>0$.
Compare this definition with the discrete case where

$$p_{X|Y}(x|y)=\frac{p_{XY}(x,y)}{p_Y(y)}.$$

Example 44.1
Suppose X and Y have the following joint density

$$f_{XY}(x,y) = \begin{cases} \frac{1}{2} & |X|+|Y|<1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the marginal distribution of X.
(b) Find the conditional distribution of Y given $X = \frac{1}{2}$.

Solution.
(a) Clearly, X only takes values in $(-1,1)$. So $f_X(x) = 0$ if $|x| \geq 1$. Let $-1 < x < 1$,

$$f_X(x) = \int_{-\infty}^{\infty} \frac{1}{2} dy = \int_{-1+|x|}^{1-|x|} \frac{1}{2} dy = 1 - |x|.$$

(b) The conditional density of Y given $X = \frac{1}{2}$ is then given by

$$f_{Y|X}(y|x) = \frac{f(\frac{1}{2},y)}{f_X(\frac{1}{2})} = \begin{cases} 1 & -\frac{1}{2} < y < \frac{1}{2} \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $f_{Y|X}$ follows a uniform distribution on the interval $\left(-\frac{1}{2},\frac{1}{2}\right)$ ■

Example 44.2
Suppose that X is uniformly distributed on the interval $[0,1]$ and that, given $X = x$, Y is uniformly distributed on the interval $[1-x,1]$.
(a) Determine the joint density $f_{XY}(x,y)$.
(b) Find the probability $\Pr(Y \geq \frac{1}{2})$.

Solution.
Since X is uniformly distributed on $[0,1]$, we have $f_X(x) = 1, 0 \leq x \leq 1$. Similarly, since, given $X = x$, Y is uniformly distributed on $[1-x,1]$, the conditional density of Y given $X = x$ is $\frac{1}{1-(1-x)} = \frac{1}{x}$ on the interval $[1-x,1]$; i.e., $f_{Y|X}(y|x) = \frac{1}{x}, 1-x \leq y \leq 1$ for $0 \leq x \leq 1$. Thus

$$f_{XY}(x,y) = f_X(x)f_{Y|X}(y|x) = \frac{1}{x}, 0 < x < 1, 1-x < y < 1.$$

(b) Using Figure 44.1 we find

$$\begin{aligned}\Pr(Y \geq \frac{1}{2}) =& \int_0^{\frac{1}{2}} \int_{1-x}^{1} \frac{1}{x} dydx + \int_{\frac{1}{2}}^{1} \int_{\frac{1}{2}}^{1} \frac{1}{x} dydx \\ =& \int_0^{\frac{1}{2}} \frac{1-(1-x)}{x} dx + \int_{\frac{1}{2}}^{1} \frac{1/2}{x} dx \\ =& \frac{1+\ln 2}{2} \blacksquare\end{aligned}$$

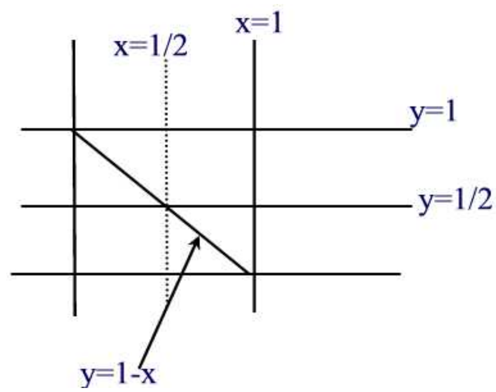


Figure 44.1

Note that

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx = \int_{-\infty}^{\infty} \frac{f_{XY}(x,y)}{f_Y(y)} dx = \frac{f_Y(y)}{f_Y(y)} = 1.$$

The **conditional cumulative distribution function** of X given $Y = y$ is defined by

$$F_{X|Y}(x|y) = \Pr(X \leq x | Y = y) = \int_{-\infty}^{x} f_{X|Y}(t|y)dt.$$

From this definition, it follows

$$f_{X|Y}(x|y) = \frac{\partial}{\partial x} F_{X|Y}(x|y).$$

Example 44.3

The joint density of X and Y is given by

$$f_{XY}(x,y) = \begin{cases} \frac{15}{2} x(2-x-y) & 0 \leq x, y \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Compute the conditional density of X, given that $Y = y$ for $0 \leq y \leq 1$.

Solution.
The marginal density function of Y is

$$f_Y(y) = \int_0^1 \frac{15}{2}x(2-x-y)dx = \frac{15}{2}\left(\frac{2}{3} - \frac{y}{2}\right).$$

Thus,

$$\begin{aligned} f_{X|Y}(x|y) =& \frac{f_{XY}(x,y)}{f_Y(y)} \\ =& \frac{x(2-x-y)}{\frac{2}{3}-\frac{y}{2}} \\ =& \frac{6x(2-x-y)}{4-3y} \end{aligned} \blacksquare$$

Example 44.4
The joint density function of X and Y is given by

$$f_{XY}(x,y) = \begin{cases} \frac{e^{-\frac{x}{y}}e^{-y}}{y} & x \geq 0, y \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Compute $\Pr(X > 1|Y = y)$.

Solution.
The marginal density function of Y is

$$f_Y(y) = e^{-y}\int_0^\infty \frac{1}{y}e^{-\frac{x}{y}}dx = -e^{-y}\left[-e^{-\frac{x}{y}}\right]_0^\infty = e^{-y}.$$

Thus,

$$\begin{aligned} f_{X|Y}(x|y) =& \frac{f_{XY}(x,y)}{f_Y(y)} \\ =& \frac{\frac{e^{-\frac{x}{y}}e^{-y}}{y}}{e^{-y}} \\ =& \frac{1}{y}e^{-\frac{x}{y}} \end{aligned}$$

Hence,

$$\begin{aligned}\Pr(X > 1|Y = y) =& \int_1^\infty \frac{1}{y}e^{-\frac{x}{y}}dx\\ =& -e^{-\frac{x}{y}}|_1^\infty = e^{-\frac{1}{y}} \blacksquare\end{aligned}$$

We end this section with the following theorem.

Theorem 44.1
Continuous random variables X and Y with $f_Y(y) > 0$ are independent if and only if

$$f_{X|Y}(x|y) = f_X(x).$$

Proof.
Suppose first that X and Y are independent. Then $f_{XY}(x,y) = f_X(x)f_Y(y)$. Thus,

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{f_X(x)f_Y(y)}{f_Y(y)} = f_X(x).$$

Conversely, suppose that $f_{X|Y}(x|y) = f_X(x)$. Then $f_{XY}(x,y) = f_{X|Y}(x|y)f_Y(y) = f_X(x)f_Y(y)$. This shows that X and Y are independent $\blacksquare$

Example 44.5
Let X and Y be two continuous random variables with joint density function

$$f_{XY}(x,y) = \begin{cases} \frac{1}{2} & 0 \le y < x \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $f_X(x), f_Y(y)$ and $f_{X|Y}(x|1)$.
(b) Are X and Y independent?

Solution.
(a) We have

$$f_X(x) = \int_0^x \frac{1}{2}dy = \frac{x}{2}, \quad 0 \le x \le 2$$

$$f_Y(y) = \int_y^2 \frac{1}{2}dx = \frac{1}{2}(2-y), \quad 0 \le y \le 2$$

and

$$f_{X|Y}(x|1) = \frac{f_{XY}(x,1)}{f_Y(1)} = \frac{\frac{1}{2}}{\frac{1}{2}} = 1, \quad 1 \le x \le 2.$$

(b) Since $f_{X|Y}(x|1) \ne f_X(x)$, X and Y are dependent $\blacksquare$

Practice Problems

Problem 44.1
Let X and Y be two random variables with joint density function

$$f_{XY}(x,y) = \begin{cases} 5x^2y & -1 \le x \le 1, 0 < y \le |x| \\ 0 & \text{otherwise.} \end{cases}$$

Find $f_{X|Y}(x|y)$, the conditional probability density function of X given $Y = y$. Sketch the graph of $f_{X|Y}(x|0.5)$.

Problem 44.2
Suppose that X and Y have joint density function

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

Find $f_{X|Y}(x|y)$, the conditional probability density function of X given $Y = y$.

Problem 44.3
Suppose that X and Y have joint density function

$$f_{XY}(x,y) = \begin{cases} \frac{3y^2}{x^3} & 0 \le y < x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find $f_{Y|X}(y|x)$, the conditional probability density function of Y given $X = x$.

Problem 44.4
The joint density function of X and Y is given by

$$f_{XY}(x,y) = \begin{cases} xe^{-x(y+1)} & x \ge 0, y \ge 0 \\ 0 & \text{otherwise.} \end{cases}$$

Find the conditional density of X given $Y = y$ and that of Y given $X = x$.

Problem 44.5
Let X and Y be continuous random variables with conditional and marginal p.d.f.'s given by

$$f_X(x) = \frac{x^3e^{-x}}{6}I_{(0,\infty)}(x)$$

and

$$f_{Y|x}(y|x) = \frac{3y^2}{x^3} I_{(0,x)}(y)$$

where $I_A(x)$ is the indicator function of A.
(a) Find the joint p.d.f. of X and Y.
(b) Find the conditional p.d.f. of X given $Y = y$.

Problem 44.6
Suppose X, Y are two continuous random variables with joint probability density function

$$f_{XY}(x,y) = \begin{cases} 12xy(1-x) & 0 < x, y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $f_{X|Y}(x|y)$. Are X and Y independent?
(b) Find $\Pr(Y < \frac{1}{2}|X > \frac{1}{2})$.

Problem 44.7
The joint probability density function of the random variables X and Y is given by

$$f_{XY}(x,y) = \begin{cases} \frac{1}{3}x - y + 1 & 1 \le x \le 2, 0 \le y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the conditional probability density function of X given $Y = y$.
(b) Find $\Pr(X < \frac{3}{2}|Y = \frac{1}{2})$.

Problem 44.8 ‡
Let X and Y be continuous random variables with joint density function

$$f_{XY}(x,y) = \begin{cases} 24xy & 0 < x < 1, 0 < y < 1 - x \\ 0 & \text{otherwise.} \end{cases}$$

Calculate $\Pr\left(Y < X | X = \frac{1}{3}\right)$.

Problem 44.9 ‡
Once a fire is reported to a fire insurance company, the company makes an initial estimate, X, of the amount it will pay to the claimant for the fire loss. When the claim is finally settled, the company pays an amount, Y, to the claimant. The company has determined that X and Y have the joint density function

$$f_{XY}(x,y) = \begin{cases} \frac{2}{x^2(x-1)} y^{-(2x-1)/(x-1)} & x > 1, y > 1 \\ 0 & \text{otherwise.} \end{cases}$$

Given that the initial claim estimated by the company is 2, determine the probability that the final settlement amount is between 1 and 3 .

Problem 44.10 ‡

A company offers a basic life insurance policy to its employees, as well as a supplemental life insurance policy. To purchase the supplemental policy, an employee must first purchase the basic policy.

Let X denote the proportion of employees who purchase the basic policy, and Y the proportion of employees who purchase the supplemental policy. Let X and Y have the joint density function $f_{XY}(x,y) = 2(x+y)$ on the region where the density is positive.

Given that 10% of the employees buy the basic policy, what is the probability that fewer than 5% buy the supplemental policy?

Problem 44.11 ‡

An auto insurance policy will pay for damage to both the policyholder's car and the other driver's car in the event that the policyholder is responsible for an accident. The size of the payment for damage to the policyholder's car, X, has a marginal density function of 1 for $0 < x < 1$. Given $X = x$, the size of the payment for damage to the other driver's car, Y, has conditional density of 1 for $x < y < x+1$.

If the policyholder is responsible for an accident, what is the probability that the payment for damage to the other driver's car will be greater than 0.5?

Problem 44.12 ‡

You are given the following information about N, the annual number of claims for a randomly selected insured:

$$\begin{aligned}
\Pr(N=0) =& \frac{1}{2}\\
\Pr(N=1) =& \frac{1}{3}\\
\Pr(N>1) =& \frac{1}{6}
\end{aligned}$$

Let S denote the total annual claim amount for an insured. When $N = 1, S$ is exponentially distributed with mean 5 . When $N > 1, S$ is exponentially distributed with mean 8 . Determine $\Pr(4 < S < 8)$.

Problem 44.13
Let Y have a uniform distribution on the interval $(0,1)$, and let the conditional distribution of X given $Y=y$ be uniform on the interval $(0,\sqrt{y})$. What is the marginal density function of X for $0<x<1$?

Problem 44.14 ‡
The distribution of Y, given X, is uniform on the interval $[0,X]$. The marginal density of X is

$$f_X(x)=\begin{cases} 2x & \text{for } 0<x<1 \\ 0 & \text{otherwise.}\end{cases}$$

Determine the conditional density of X, given $Y=y>0$.

Problem 44.15
Suppose that X has a continuous distribution with p.d.f. $f_X(x)=2x$ on $(0,1)$ and 0 elsewhere. Suppose that Y is a continuous random variable such that the conditional distribution of Y given $X=x$ is uniform on the interval $(0,x)$. Find the mean and variance of Y.

Problem 44.16 ‡
An insurance policy is written to cover a loss X where X has density function

$$f_X(x)=\begin{cases} \frac{3}{8}x^2 & 0\le x\le 2 \\ 0 & \text{otherwise.}\end{cases}$$

The time T (in hours) to process a claim of size x, where $0\le x\le 2$, is uniformly distributed on the interval from x to $2x$.
Calculate the probability that a randomly chosen claim on this policy is processed in three hours or more.

45 Joint Probability Distributions of Functions of Random Variables

Theorem 38.1 provided a result for finding the pdf of a function of one random variable: if $Y = g(X)$ is a function of the random variable X, where $g(x)$ is monotone and differentiable then the pdf of Y is given by

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

An extension to functions of two random variables is given in the following theorem.

Theorem 45.1
Let X and Y be jointly continuous random variables with joint probability density function $f_{XY}(x,y)$. Let $U = g_1(X,Y)$ and $V = g_2(X,Y)$. Assume that the functions $u = g_1(x,y)$ and $v = g_2(x,y)$ can be solved uniquely for x and y. Furthermore, suppose that g_1 and g_2 have continuous partial derivatives at all points (x,y) and such that the Jacobian determinant

$$J(x,y) = \begin{vmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \end{vmatrix} = \frac{\partial g_1}{\partial x}\frac{\partial g_2}{\partial y} - \frac{\partial g_1}{\partial y}\frac{\partial g_2}{\partial x} \neq 0$$

for all x and y. Then the random variables U and V are continuous random variables with joint density function given by

$$f_{UV}(u,v) = f_{XY}(x(u,v), y(u,v))|J(x(u,v), y(u,v))|^{-1}$$

Proof.
We first remind the reader about the change of variable formula for a double integral. Suppose $x = x(u,v)$ and $y = y(u,v)$ are two differentiable functions of u and v. We assume that the functions x and y take a point in the $uv-$plane to exactly one point in the $xy-$plane.
Let us see what happens to a small rectangle T in the $uv-$plane with sides of lengths Δu and Δv as shown in Figure 45.1.

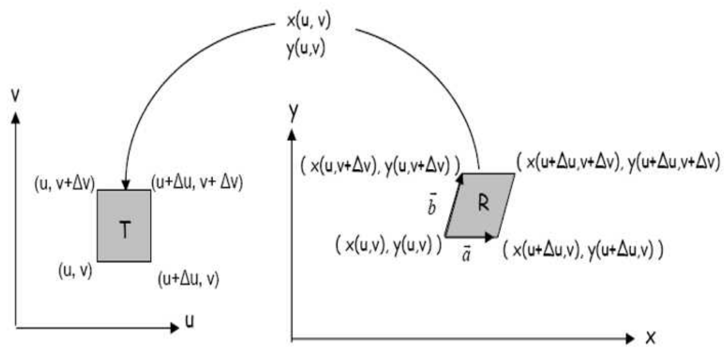


Figure 45.1

Since the side-lengths are small, by local linearity each side of the rectangle in the $uv-$plane is transformed into a line segment in the $xy-$plane. The result is that the rectangle in the $uv-$plane is transformed into a parallelogram R in the $xy-$plane with sides in vector form

$$\vec{a} = [x(u+\Delta u, v) - x(u,v)]\vec{i} + [y(u+\Delta u, v) - y(u,v)]\vec{j} \approx \frac{\partial x}{\partial u}\Delta u\vec{i} + \frac{\partial y}{\partial u}\Delta u\vec{j}$$

and

$$\vec{b} = [x(u, v+\Delta v) - x(u,v)]\vec{i} + [y(u, v+\Delta v) - y(u,v)]\vec{j} \approx \frac{\partial x}{\partial v}\Delta v\vec{i} + \frac{\partial y}{\partial v}\Delta v\vec{j}$$

Now, the area of R is

$$\text{Area R} \approx ||\vec{a}\times\vec{b}|| = \left|\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\right|\Delta u\Delta v$$

Using determinant notation, we define the **Jacobian**, $\frac{\partial(x,y)}{\partial(u,v)}$, as follows

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

Thus, we can write

$$\text{Area R} \approx \left|\frac{\partial(x,y)}{\partial(u,v)}\right|\Delta u\Delta v$$

Now, suppose we are integrating $f(x, y)$ over a region R. Partition R into mn small parallelograms. Then using Riemann sums we can write

$$\begin{aligned}\int_R f(x,y)dxdy \approx &\sum_{j=1}^{m}\sum_{i=1}^{n} f(x_{ij}, y_{ij}) \cdot \text{Area of } \mathrm{R_{ij}} \\ \approx &\sum_{j=1}^{m}\sum_{i=1}^{n} f(x(u_{ij}, v_{ij}), y(u_{ij}, v_{ij})) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \Delta u \Delta v\end{aligned}$$

where (x_{ij}, y_{ij}) in R_{ij} corresponds to a point (u_{ij}, v_{ij}) in T_{ij}. Now, letting $m, n \to \infty$ to otbain

$$\int_R f(x,y)dxdy = \int_T f(x(u,v), y(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| dudv.$$

The result of the theorem follows from the fact that if a region R in the $xy-$plane maps into the region T in the $uv-$plane then we must have

$$\begin{aligned}\Pr((X,Y) \in R) = &\int\int_R f_{XY}(x,y)dxdy \\ = &\int\int_T f_{XY}(x(u,v), y(u,v))|J(x(u,v), y(u,v))|^{-1} dudv \\ = &\Pr((U,V) \in T) \blacksquare\end{aligned}$$

Example 45.1
Let X and Y be jointly continuous random variables with density function $f_{XY}(x, y)$. Let $U = X + Y$ and $V = X - Y$. Find the joint density function of U and V.

Solution.
Let $u = g_1(x, y) = x+y$ and $v = g_2(x, y) = x-y$. Then $x = \frac{u+v}{2}$ and $y = \frac{u-v}{2}$. Moreover

$$J(x,y) = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2$$

Thus,

$$f_{UV}(u,v) = \frac{1}{2} f_{XY}\left(\frac{u+v}{2}, \frac{u-v}{2}\right) \blacksquare$$

Example 45.2
Let X and Y be jointly continuous random variables with density function $f_{XY}(x,y) = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$. Let $U = X+Y$ and $V = X-Y$. Find the joint density function of U and V.

Solution.
Since $J(x,y) = -2$ we have

$$f_{UV}(u,v) = \frac{1}{4\pi}e^{-\frac{(\frac{u+v}{2})^2+(\frac{u-v}{2})^2}{2}} = \frac{1}{4\pi}e^{-\frac{u^2+v^2}{4}} \blacksquare$$

Example 45.3
Suppose that X and Y have joint density function given by

$$f_{XY}(x,y) = \begin{cases} 4xy & 0<x<1,\ 0<y<1 \\ 0 & \text{otherwise} \end{cases}$$

Let $U = \frac{X}{Y}$ and $V = XY$.
(a) Find the joint density function of U and V.
(b) Find the marginal density of U and V.
(c) Are U and V independent?

Solution.
(a) Now, if $u = g_1(x,y) = \frac{x}{y}$ and $v = g_2(x,y) = xy$ then solving for x and y we find $x = \sqrt{uv}$ and $y = \sqrt{\frac{v}{u}}$. Moreover,

$$J(x,y) = \begin{vmatrix} \frac{1}{y} & -\frac{x}{y^2} \\ y & x \end{vmatrix} = \frac{2x}{y}$$

By Theorem 45.1, we find

$$f_{UV}(u,v) = \frac{1}{2u}f_{XY}(\sqrt{uv}, \sqrt{\frac{v}{u}}) = \frac{2v}{u},\ 0<uv<1,\ 0<\frac{v}{u}<1$$

and 0 otherwise. The region where f_{UV} is defined is shown in Figure 45.2.
(b) The marginal density of U is

$$f_U(u) = \int_0^u \frac{2v}{u}dv = u, \quad 0<u\le 1$$
$$f_U(u) = \int_0^{\frac{1}{u}} \frac{2v}{u}dv = \frac{1}{u^3}, \quad u>1$$

and the marginal density of V is

$$f_V(v) = \int_0^\infty f_{UV}(u,v)du = \int_v^{\frac{1}{v}} \frac{2v}{u} du = -4v \ln v, \quad 0 < v < 1$$

(c) Since $f_{UV}(u,v) \neq f_U(u)f_V(v)$, U and V are dependent ∎

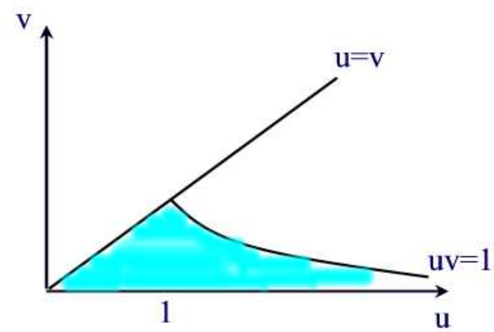


Figure 45.2

Practice Problems

Problem 45.1
Let X and Y be two random variables with joint pdf f_{XY}. Let $Z = aX + bY$ and $W = cX + dY$ where $ad - bc \neq 0$. Find the joint probability density function of Z and W.

Problem 45.2
Let X_1 and X_2 be two independent exponential random variables each having parameter λ. Find the joint density function of $Y_1 = X_1 + X_2$ and $Y_2 = e^{X_2}$.

Problem 45.3
Let X and Y be random variables with joint pdf $f_{XY}(x, y)$. Let $R = \sqrt{X^2 + Y^2}$ and $\Phi = \tan^{-1}\left(\frac{Y}{X}\right)$ with $-\pi < \Phi \leq \pi$. Find $f_{R\Phi}(r, \phi)$.

Problem 45.4
Let X and Y be two random variables with joint pdf $f_{XY}(x, y)$. Let $Z = g(X, Y) = \sqrt{X^2 + Y^2}$ and $W = \frac{Y}{X}$. Find $f_{ZW}(z, w)$.

Problem 45.5
If X and Y are independent gamma random variables with parameters (α, λ) and (β, λ) respectively, compute the joint density of $U = X + Y$ and $V = \frac{X}{X+Y}$.

Problem 45.6
Let X_1 and X_2 be two continuous random variables with joint density function

$$f_{X_1X_2}(x_1, x_2) = \begin{cases} e^{-(x_1+x_2)} & x_1 \geq 0, x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Let $Y_1 = X_1 + X_2$ and $Y_2 = \frac{X_1}{X_1+X_2}$. Find the joint density function of Y_1 and Y_2.

Problem 45.7
Let X_1 and X_2 be two independent normal random variables with parameters (0,1) and (0,4) respectively. Let $Y_1 = 2X_1 + X_2$ and $Y_2 = X_1 - 3X_2$. Find $f_{Y_1Y_2}(y_1, y_2)$.

Problem 45.8
Let X be a uniform random variable on $(0, 2\pi)$ and Y an exponential random variable with $\lambda = 1$ and independent of X. Show that

$$U = \sqrt{2Y}\cos X \text{ and } V = \sqrt{2Y}\sin X$$

are independent standard normal random variables

Problem 45.9
Let X and Y be two random variables with joint density function f_{XY}. Compute the pdf of $U = X + Y$. What is the pdf in the case X and Y are independent? Hint: let $V = Y$.

Problem 45.10
Let X and Y be two random variables with joint density function f_{XY}. Compute the pdf of $U = Y - X$.

Problem 45.11
Let X and Y be two random variables with joint density function f_{XY}. Compute the pdf of $U = XY$. Hint: let $V = X$.

Problem 45.12
Let X and Y be two independent exponential distributions with mean 1. Find the distribution of $\frac{X}{Y}$.

Properties of Expectation

We have seen that the expected value of a random variable is a weighted average of the possible values of X and also is the center of the distribution of the variable. Recall that the expected value of a discrete random variable X with probability mass function $p(x)$ is defined by

$$E(X) = \sum_x xp(x)$$

provided that the sum is finite.
For a continuous random variable X with probability density function $f(x)$, the expected value is given by

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

provided that the improper integral is convergent.
In this chapter we develop and exploit properties of expected values.

46 Expected Value of a Function of Two Random Variables

In this section, we learn some equalities and inequalities about the expectation of random variables. Our goals are to become comfortable with the expectation operator and learn about some useful properties.
First, we introduce the definition of expectation of a function of two random variables: Suppose that X and Y are two random variables taking values in S_X and S_Y respectively. For a function $g : S_X \times S_Y \to \mathbb{R}$ the expected value of $g(X, Y)$ is

$$E(g(X, Y) = \sum_{x \in S_X} \sum_{y \in S_Y} g(x, y)p_{XY}(x, y).$$

if X and Y are discrete with joint probability mass function $p_{XY}(x,y)$ and

$$E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{XY}(x,y)dxdy$$

if X and Y are continuous with joint probability density function $f_{XY}(x,y)$.

Example 46.1
Let X and Y be two discrete random variables with joint probability mass function:

$$p_{XY}(1,1) = \tfrac{1}{3}, p_{XY}(1,2) = \tfrac{1}{8}, p_{XY}(2,1) = \tfrac{1}{2}, p_{XY}(2,2) = \tfrac{1}{24}$$

Find the expected value of $g(X,Y) = XY$.

Solution.
The expected value of the function $g(X,Y) = XY$ is calculated as follows:

$$\begin{aligned}
E(g(X,Y)) =&E(XY) = \sum_{x=1}^{2}\sum_{y=1}^{2} xyp_{XY}(x,y)\\
=&(1)(1)(\frac{1}{3}) + (1)(2)(\frac{1}{8}) + (2)(1)(\frac{1}{2}) + (2)(2)(\frac{1}{24})\\
=&\frac{7}{4} \blacksquare
\end{aligned}$$

An important application of the above definition is the following result.

Proposition 46.1
The expected value of the sum/difference of two random variables is equal to the sum/difference of their expectations. That is,

$$E(X+Y) = E(X) + E(Y)$$

and

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

Proof.
We prove the result for discrete random variables X and Y with joint probability mass function $p_{XY}(x,y)$. Letting $g(X,Y) = X \pm Y$ we have

$$\begin{aligned}
E(X \pm Y) =& \sum_x \sum_y (x \pm y) p_{XY}(x,y) \\
=& \sum_x \sum_y x p_{XY}(x,y) \pm \sum_x \sum_y y p_{XY}(x,y) \\
=& \sum_x x \sum_y p_{XY}(x,y) \pm \sum_y y \sum_x p_{XY}(x,y) \\
=& \sum_x x p_X(x) \pm \sum_y y p_Y(y) \\
=& E(X) \pm E(Y)
\end{aligned}$$

A similar proof holds for the continuous case where you just need to replace the sums by improper integrals and the joint probability mass function by the joint probability density function ■

Using mathematical induction one can easily extend the previous result to

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n), \quad E(X_i) < \infty.$$

Example 46.2
A group of N business executives throw their business cards into a jar. The cards are mixed, and each person randomly selects one. Find the expected number of people that select their own card.

Solution.
Let X = the number of people who select their own card. For $1 \le i \le N$ let

$$X_i = \begin{cases} 1 & \text{if the } i^{\text{th}} \text{ person chooses his own card} \\ 0 & \text{otherwise} \end{cases}$$

Then $E(X_i) = \Pr(X_i = 1) = \frac{1}{N}$ and

$$X = X_1 + X_2 + \cdots + X_N.$$

Hence,

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_N) = \left(\frac{1}{N}\right) N = 1 \blacksquare$$

Example 46.3 *(***Sample Mean)**
Let $X_1, X_2, \cdots, X_n$ be a sequence of independent and identically distributed random variables, each having a mean μ and variance σ^2. Define a new random variable by

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

We call $\overline{X}$ the **sample mean**. Find $E(\overline{X})$.

Solution.
The expected value of $\overline{X}$ is

$$E(\overline{X}) = E\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right] = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu.$$

Because of this result, when the distribution mean μ is unknown, the sample mean is often used in statisitcs to estimate it ■

The following property is known as the monotonicity property of the expected value.

Proposition 46.2
If X is a nonnegative random variable then $E(X) \geq 0$. Thus, if X and Y are two random variables such that $X \geq Y$ then $E(X) \geq E(Y)$.

Proof.
We prove the result for the continuous case. We have

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x)dx \\ &= \int_{0}^{\infty} xf(x)dx \geq 0 \end{aligned}$$

since $f(x) \geq 0$ so the integrand is nonnegative. Now, if $X \geq Y$ then $X - Y \geq 0$ so that by the previous proposition we can write $E(X) - E(Y) = E(X - Y) \geq 0$ ■

As a direct application of the monotonicity property we have

Proposition 46.3 *(Boole's Inequality)*
For any events $A_1, A_2, \cdots, A_n$ we have

$$\Pr\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \Pr(A_i).$$

Proof.
For $i = 1, \cdots, n$ define

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

Let

$$X = \sum_{i=1}^{n} X_i$$

so X denotes the number of the events A_i that occur. Also, let

$$Y = \begin{cases} 1 & \text{if } X \geq 1 \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

so Y is equal to 1 if at least one of the A_i occurs and 0 otherwise. Clearly, $X \geq Y$ so that $E(X) \geq E(Y)$. But

$$E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} P(A_i)$$

and

$$E(Y) = \Pr\{ \text{ at least one of the } A_i \text{ occur } \} = \Pr\left(\bigcup_{i=1}^{n} A_i\right).$$

Thus, the result follows. Note that for any set A we have

$$E(I_A) = \int I_A(x) f(x) dx = \int_A f(x) dx = \Pr(A) \blacksquare$$

Proposition 46.4
If X is a random variable with range $[a, b]$ then $a \leq E(X) \leq b$.

Proof.
Let $Y = X - a \geq 0$. Then $E(Y) \geq 0$. But $E(Y) = E(X) - E(a) = E(X) - a \geq 0$. Thus, $E(X) \geq a$. Similarly, let $Z = b - X \geq 0$. Then $E(Z) = b - E(X) \geq 0$ or $E(X) \leq b$ ■

We have determined that the expectation of a sum is the sum of the expectations. The same is not always true for products: in general, the expectation of a product need not equal the product of the expectations. But it is true in an important special case, namely, when the random variables are independent.

Proposition 46.5
If X and Y are independent random variables then for any function h and g we have

$$E(g(X)h(Y)) = E(g(X))E(h(Y))$$

In particular, $E(XY) = E(X)E(Y)$.

Proof.
We prove the result for the continuous case. The proof of the discrete case is similar. Let X and Y be two independent random variables with joint density function $f_{XY}(x,y)$. Then

$$\begin{aligned}
E(g(X)h(Y)) =& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_{XY}(x,y)dxdy \\
=& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dxdy \\
=& \left(\int_{-\infty}^{\infty} h(y)f_Y(y)dy\right)\left(\int_{-\infty}^{\infty} g(x)f_X(x)dx\right) \\
=& E(h(Y))E(g(X)) \blacksquare
\end{aligned}$$

We next give a simple example to show that the expected values need not multiply if the random variables are not independent.

Example 46.4
Consider a single toss of a coin. We define the random variable X to be 1 if heads turns up and 0 if tails turns up, and we set $Y = 1 - X$. Thus X and Y are dependent. Show that $E(XY) \neq E(X)E(Y)$.

Solution.
Clearly, $E(X) = E(Y) = \frac{1}{2}$. But $XY = 0$ so that $E(XY) = 0 \neq E(X)E(Y)$ ■

Example 46.5
Suppose a box contains 10 green, 10 red and 10 black balls. We draw 10 balls from the box by sampling with replacement. Let X be the number of green balls, and Y be the number of black balls in the sample.
(a) Find $E(XY)$.
(b) Are X and Y independent? Explain.

Solution.
First we note that X and Y are binomial with $n = 10$ and $p = \frac{1}{3}$.
(a) Let X_i be 1 if we get a green ball on the i^{th} draw and 0 otherwise, and Y_j be the event that in j^{th} draw we got a black ball. Trivially, X_i and Y_j are independent if $1 \leq i \neq j \leq 10$. Moreover, $X_iY_i = 0$ for all $1 \leq i \leq 10$. Since $X = X_1 + X_2 + \cdots X_{10}$ and $Y = Y_1 + Y_2 + \cdots Y_{10}$ we have

$$XY = \sum \sum_{1 \leq i \neq j \leq 10} X_iY_j.$$

Hence,

$$E(XY) = \sum \sum_{1 \leq i \neq j \leq 10} E(X_iY_j) = \sum \sum_{1 \leq i \neq j \leq 10} E(X_i)E(Y_j) = 90 \times \frac{1}{3} \times \frac{1}{3} = 10.$$

(b) Since $E(X) = E(Y) = \frac{10}{3}$, we have $E(XY) \neq E(X)E(Y)$ so X and Y are dependent ■

The following inequality will be of importance in the next section

Proposition 46.6 *(Markov's Inequality)*
If $X \geq 0$ and $c > 0$ then $\Pr(X \geq c) \leq \frac{E(X)}{c}$.

Proof.
Let $c > 0$. Define

$$I = \begin{cases} 1 & \text{if } X \geq c \\ 0 & \text{otherwise} \end{cases}$$

Since $X \geq 0$, we have $I \leq \frac{X}{c}$. Taking expectations of both side we find $E(I) \leq \frac{E(X)}{c}$. Now the result follows since $E(I) = \Pr(X \geq c)$ ■

Example 46.6
Let X be a non-negative random variable. Let a be a positive constant. Prove that $\Pr(X \geq a) \leq \frac{E(e^{tX})}{e^{ta}}$ for all $t \geq 0$.

Solution.
Applying Markov's inequality we find

$$\Pr(X \geq a) = \Pr(tX \geq ta) = \Pr(e^{tX} \geq e^{ta}) \leq \frac{E(e^{tX})}{e^{ta}} \blacksquare$$

As an important application of the previous result we have

Proposition 46.7
If $X \geq 0$ and $E(X) = 0$ then $\Pr(X = 0) = 1$.

Proof.
Since $E(X) = 0$, by the previous result we find $\Pr(X \geq c) = 0$ for all $c > 0$. But

$$\Pr(X > 0) = \Pr\left(\bigcup_{n=1}^{\infty}(X > \frac{1}{n})\right) \leq \sum_{n=1}^{\infty}\Pr(X > \frac{1}{n}) = 0.$$

Hence, $P(X > 0) = 0$. Since $X \geq 0$, we have $1 = \Pr(X \geq 0) = \Pr(X = 0) + \Pr(X > 0) = \Pr(X = 0)$ $\blacksquare$

Corollary 46.1
Let X be a random variable. If $Var(X) = 0$, then $\Pr(X = E(X)) = 1$.

Proof.
Suppose that $Var(X) = 0$. Since $(X - E(X))^2 \geq 0$ and $Var(X) = E((X - E(X))^2)$, by the previous result we have $P(X - E(X) = 0) = 1$. That is, $\Pr(X = E(X)) = 1$ $\blacksquare$

Example 46.7 *(expected value of a Binomial Random Variable)*
Let X be a binomial random variable with parameters (n, p). Find $E(X)$.

Solution.
We have that X is the number of successes in n trials. For $1 \leq i \leq n$ let X_i denote the number of successes in the ith trial. Then $E(X_i) = 0(1-p)+1p = p$. Since $X = X_1 + X_2 + \cdots + X_n$, we find $E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = np$ $\blacksquare$

Practice Problems

Problem 46.1
Let X and Y be independent random variables, both being equally likely to be any of the numbers $1, 2, \cdots, m$. Find $E(|X - Y|)$.

Problem 46.2
Let X and Y be random variables with joint pdf

$$f_{XY}(x, y) = \begin{cases} 1 & 0 < x < 1, x < y < x + 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $E(XY)$.

Problem 46.3
Let X and Y be two independent uniformly distributed random variables in [0,1]. Find $E(|X - Y|)$.

Problem 46.4
Let X and Y be continuous random variables with joint pdf

$$f_{XY}(x, y) = \begin{cases} 2(x + y) & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $E(X^2Y)$ and $E(X^2 + Y^2)$.

Problem 46.5
Suppose that $E(X) = 5$ and $E(Y) = -2$. Find $E(3X + 4Y - 7)$.

Problem 46.6
Suppose that X and Y are independent, and that $E(X) = 5$, $E(Y) = -2$. Find $E[(3X - 4)(2Y + 7)]$.

Problem 46.7
Let X and Y be two independent random variables that are uniformly distributed on the interval $(0, L)$. Find $E(|X - Y|)$.

Problem 46.8
Ten married couples are to be seated at five different tables, with four people at each table. Assume random seating, what is the expected number of married couples that are seated at the same table?

Problem 46.9
John and Katie randomly, and independently, choose 3 out of 10 objects. Find the expected number of objects
(a) chosen by both individuals.
(b) not chosen by either individual.
(c) chosen exactly by one of the two.

Problem 46.10
If $E(X) = 1$ and Var(X) = 5 find
(a) $E[(2+X)^2]$
(b) $\mathrm{Var}(4+3X)$

Problem 46.11 ‡
Let T_1 be the time between a car accident and reporting a claim to the insurance company. Let T_2 be the time between the report of the claim and payment of the claim. The joint density function of T_1 and T_2, $f(t_1, t_2)$, is constant over the region $0 < t_1 < 6, 0 < t_2 < 6, t_1 + t_2 < 10$, and zero otherwise.
Determine $E[T_1+T_2]$, the expected time between a car accident and payment of the claim.

Problem 46.12 ‡
Let T_1 and T_2 represent the lifetimes in hours of two linked components in an electronic device. The joint density function for T_1 and T_2 is uniform over the region defined by $0 \le t_1 \le t_2 \le L$, where L is a positive constant.
Determine the expected value of the sum of the squares of T_1 and T_2.

Problem 46.13
Let X and Y be two independent random variables with $\mu_X = 1, \mu_Y = -1, \sigma_X^2 = \frac{1}{2}$, and $\sigma_Y^2 = 2$. Compute $E[(X+1)^2(Y-1)^2]$.

Problem 46.14 ‡
A machine consists of two components, whose lifetimes have the joint density function

$$f(x,y) = \begin{cases} \frac{1}{50} & \text{for } x > 0, y > 0, x + y < 10 \\ 0 & \text{otherwise.} \end{cases}$$

The machine operates until both components fail. Calculate the expected operational time of the machine.

47 Covariance, Variance of Sums, and Correlations

So far, We have discussed the absence or presence of a relationship between two random variables, i.e. independence or dependence. But if there is in fact a relationship, the relationship may be either weak or strong. For example, if X is the weight of a sample of water and Y is the volume of the sample of water then there is a strong relationship between X and Y. On the other hand, if X is the weight of a person and Y denotes the same person's height then there is a relationship between X and Y but not as strong as in the previous example.
We would like a measure that can quantify this difference in the strength of a relationship between two random variables.
The **covariance** between X and Y is defined by

$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))].$$

An alternative expression that is sometimes more convenient is

$$\begin{aligned}Cov(X,Y) =&E(XY - E(X)Y - XE(Y) + E(X)E(Y))\\ =&E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)\\ =&E(XY) - E(X)E(Y).\end{aligned}$$

Recall that for independent X, Y we have $E(XY) = E(X)E(Y)$ and so $Cov(X,Y) = 0$. However, the converse statement is false as there exists random variables that have covariance 0 but are dependent. For example, let X be a random variable such that

$$\Pr(X = 0) = \Pr(X = 1) = \Pr(X = -1) = \frac{1}{3}$$

and define

$$Y = \begin{cases} 0 & \text{if } X \neq 0 \\ 1 & \text{otherwise.} \end{cases}$$

Thus, Y depends on X.
Clearly, $XY = 0$ so that $E(XY) = 0$. Also,

$$E(X) = (0 + 1 - 1)\frac{1}{3} = 0$$

and thus

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 0.$$

Useful facts are collected in the next result.

Theorem 47.1
(a) $Cov(X,Y) = Cov(Y,X)$ (Symmetry)
(b) $Cov(X,X) = Var(X)$
(c) $Cov(aX,Y) = aCov(X,Y)$
(d) $Cov\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m Cov(X_i, Y_j)$

Proof.
(a) $Cov(X,Y) = E(XY)-E(X)E(Y) = E(YX)-E(Y)E(X) = Cov(Y,X)$.
(b) $Cov(X,X) = E(X^2) - (E(X))^2 = Var(X)$.
(c) $Cov(aX,Y) = E(aXY) - E(aX)E(Y) = aE(XY) - aE(X)E(Y) = a(E(XY) - E(X)E(Y)) = aCov(X,Y)$.
(d) First note that $E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E(X_i)$ and $E\left[\sum_{j=1}^m Y_j\right] = \sum_{j=1}^m E(Y_j)$. Then

$$\begin{aligned}
Cov\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) =& E\left[\left(\sum_{i=1}^n X_i - \sum_{i=1}^n E(X_i)\right)\left(\sum_{j=1}^m Y_j - \sum_{j=1}^m E(Y_j)\right)\right]\\
=& E\left[\sum_{i=1}^n (X_i - E(X_i)) \sum_{j=1}^m (Y_j - E(Y_j))\right]\\
=& E\left[\sum_{i=1}^n \sum_{j=1}^m (X_i - E(X_i))(Y_j - E(Y_j))\right]\\
=& \sum_{i=1}^n \sum_{j=1}^m E[(X_i - E(X_i))(Y_j - E(Y_j))]\\
=& \sum_{i=1}^n \sum_{j=1}^m Cov(X_i, Y_j) \blacksquare
\end{aligned}$$

Example 47.1
Given that $E(X) = 5, E(X^2) = 27.4, E(Y) = 7, E(Y^2) = 51.4$ and $\text{Var}(X + Y) = 8$, find $\text{Cov}(X + Y, X + 1.2Y)$.

Solution.
By definition,

$$\mathrm{Cov}(X+Y, X+1.2Y) = E((X+Y)(X+1.2Y)) - E(X+Y)E(X+1.2Y)$$

Using the properties of expectation and the given data, we get

$$\begin{aligned}
E(X+Y)E(X+1.2Y) =&(E(X)+E(Y))(E(X)+1.2E(Y))\\
=&(5+7)(5+(1.2)\cdot 7) = 160.8\\
E((X+Y)(X+1.2Y)) =&E(X^2)+2.2E(XY)+1.2E(Y^2)\\
=&27.4+2.2E(XY)+(1.2)(51.4)\\
=&2.2E(XY)+89.08
\end{aligned}$$

Thus,

$$\mathrm{Cov}(X+Y, X+1.2Y) = 2.2E(XY)+89.08-160.8 = 2.2E(XY)-71.72$$

To complete the calculation, it remains to find $E(XY)$. To this end we make use of the still unused relation $\mathrm{Var}(X+Y)=8$

$$\begin{aligned}
8 =&\mathrm{Var}(X+Y) = E((X+Y)^2) - (E(X+Y))^2\\
=&E(X^2)+2E(XY)+E(Y^2)-(E(X)+E(Y))^2\\
=&27.4+2E(XY)+51.4-(5+7)^2 = 2E(XY)-65.2
\end{aligned}$$

so $E(XY)=36.6$. Substituting this above gives $\mathrm{Cov}(X+Y, X+1.2Y) = (2.2)(36.6)-71.72 = 8.8$ ■

Example 47.2
Given: $E(X)=10, \mathrm{Var}(X)=25, E(Y)=50, \mathrm{Var}(Y)=100, E(Z)=6, \mathrm{Var}(Z)=4, \mathrm{Cov}(X,Y)=10$, and $\mathrm{Cov}(X,Z)=3.5$. Let $Z=X+cY$. Find c if $\mathrm{Cov}(X,Z)=3.5$.

Solution.
We have

$$\begin{aligned}
\mathrm{Cov}(X,Z) =&\mathrm{Cov}(X, X+cY) = \mathrm{Cov}(X,X)+c\mathrm{Cov}(X,Y)\\
=&\mathrm{Var}(X)+c\mathrm{Cov}(X,Y) = 25+c(10) = 3.5
\end{aligned}$$

Solving for c we find $c=-2.15$ ■

Using (b) and (d) in the previous theorem with $Y_j = X_j,\ j = 1, 2, \cdots, n$ we find

$$\begin{aligned} Var\left(\sum_{i=1}^n X_i\right) =& Cov\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i\right) \\ =& \sum_{i=1}^n \sum_{i=1}^n Cov(X_i, X_j) \\ =& \sum_{i=1}^n Var(X_i) + \quad \sum_{i \neq j} \sum Cov(X_i, X_j) \end{aligned}$$

Since each pair of indices $i \neq j$ appears twice in the double summation, the above reduces to

$$Var\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n Var(X_i) + 2 \quad \sum_{i<j} \sum Cov(X_i, X_j).$$

In particular, if $X_1, X_2, \cdots, X_n$ are pairwise independent then

$$Var\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n Var(X_i).$$

Example 47.3 ‡

The profit for a new product is given by $Z = 3X - Y - 5$, where X and Y are independent random variables with $\text{Var}(X) = 1$ and $\text{Var}(Y) = 2$. What is the variance of Z?

Solution.

Using the properties of a variance, and independence, we get

$$\begin{aligned} \text{Var}(Z) =& \text{Var}(3X - Y - 5) = \text{Var}(3X - Y) \\ =& \text{Var}(3X) + \text{Var}(-Y) = 9\text{Var}(X) + \text{Var}(Y) = 11 \ \blacksquare \end{aligned}$$

Example 47.4

A salesperson salary consists of two parts a commission, X, and a fixed income Y. so that the total salary is $X+Y$. Suppose that $\text{Var}(X) = 5,000$, $\text{Var}(Y) = 10,000$, and $\text{Var}(X+Y) = 17,000$.

If X is increased by a flat amount of 100, and Y is increased by 10%, what is the variance of the total salary after these increases?

Solution.
We need to compute $\mathrm{Var}(X + 100 + 1.1Y)$. Since adding constants does not change the variance, this is the same as $\mathrm{Var}(X + 1.1Y)$, which expands as follows:

$$\begin{aligned}\mathrm{Var}(X + 1.1Y) =&\mathrm{Var}(X) + \mathrm{Var}(1.1Y) + 2\mathrm{Cov}(X, 1.1Y)\\ =&\mathrm{Var}(X) + 1.21\mathrm{Var}(Y) + 2(1.1)\mathrm{Cov}(X, Y)\end{aligned}$$

We are given that $\mathrm{Var}(X) = 5,000, \mathrm{Var}(Y) = 10,000$, so the only remaining unknown quantity is $\mathrm{Cov}(X, Y)$, which can be computed via the general formula for $\mathrm{Var}(X + Y)$:

$$\begin{aligned}\mathrm{Cov}(X, Y) =&\frac{1}{2}(\mathrm{Var}(X + Y) - \mathrm{Var}(X) - \mathrm{Var}(Y))\\ =&\frac{1}{2}(17,000 - 5,000 - 10,000) = 1,000\end{aligned}$$

Substituting this into the above formula, we get the answer:

$$\mathrm{Var}(X + 1.1Y) = 5,000 + 1.21(10,000) + 2(1.1)(1,000) = 19,300 \blacksquare$$

Example 47.5
Let $\overline{X}$ be the sample mean of n independent random variables $X_1, X_2, \cdots, X_n$. Find $Var(\overline{X})$.

Solution.
By independence we have

$$Var(\overline{X}) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) \blacksquare$$

The following result is known as the **Cauchy Schwartz inequality**.

Theorem 47.2
Let X and Y be two random variables. Then

$$Cov(X, Y)^2 \leq Var(X)Var(Y)$$

with equality if and only if X and Y are linearly related, i.e.,

$$Y = a + bX$$

for some constants a and b with $a \neq 0$.

Proof.
We first show that the inequality holds if either $\text{Var}(X) = 0$ or $\text{Var}(Y) = 0$. Suppose that $\text{Var}(X) = 0$. Then $\Pr(X = E(X)) = 1$ and so $X = E(X)$. By the definition of covariance, we have $\text{Cov}(X, Y) = 0$. Similar argument for $\text{Var}(Y) = 0$.
So assume that $\text{Var}(X) > 0$ and $\text{Var}(Y) > 0$. Let

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}.$$

We need to show that $|\rho| \leq 1$ or equivalently $-1 \leq \rho(X, Y) \leq 1$. If we let σ_X^2 and σ_Y^2 denote the variance of X and Y respectively then we have

$$\begin{aligned}
0 \leq& Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right)\\
=&\frac{Var(X)}{\sigma_X^2} + \frac{Var(Y)}{\sigma_Y^2} + \frac{2Cov(X, Y)}{\sigma_X \sigma_Y}\\
=&2[1 + \rho(X, Y)]
\end{aligned}$$

implying that $-1 \leq \rho(X, Y)$. Similarly,

$$\begin{aligned}
0 \leq& Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right)\\
=&\frac{Var(X)}{\sigma_X^2} + \frac{Var(Y)}{\sigma_Y^2} - \frac{2Cov(X, Y)}{\sigma_X \sigma_Y}\\
=&2[1 - \rho(X, Y)]
\end{aligned}$$

implying that $\rho(X, Y) \leq 1$.
Suppose now that $Cov(X, Y)^2 = Var(X)Var(Y)$. This implies that either $\rho(X, Y) = 1$ or $\rho(X, Y) = -1$. If $\rho(X, Y) = 1$ then $Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 0$. This implies that $\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y} = C$ for some constant C (See Corollary 35.4) or $Y = a + bX$ where $b = \frac{\sigma_Y}{\sigma_X} > 0$. If $\rho(X, Y) = -1$ then $Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 0$. This implies that $\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y} = C$ or $Y = a + bX$ where $b = -\frac{\sigma_Y}{\sigma_X} < 0$.
Conversely, suppose that $Y = a + bX$. Then

$$\rho(X, Y) = \frac{E(aX + bX^2) - E(X)E(a + bX)}{\sqrt{Var(X)b^2Var(X)}} = \frac{bVar(X)}{|b|Var(X)} = sign(b).$$

If $b > 0$ then $\rho(X, Y) = 1$ and if $b < 0$ then $\rho(X, Y) = -1$ ■

The **Correlation coefficient** of two random variables X and Y (with positive variance) is defined by

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}.$$

From the above theorem we have the correlation inequality

$$-1 \leq \rho \leq 1.$$

The correlation coefficient is a measure of the degree of linearity between X and Y. A value of $\rho(X, Y)$ near $+1$ or -1 indicates a high degree of linearity between X and Y, whereas a value near 0 indicates a lack of such linearity.
Correlation is a scaled version of covariance; note that the two parameters always have the same sign (positive, negative, or 0). When the sign is positive, the variables X and Y are said to be **positively correlated** and this indicates that Y tends to increase when X does; when the sign is negative, the variables are said to be **negatively correlated** and this indicates that Y tends to decrease when X increases; and when the sign is 0, the variables are said to be **uncorrelated.**
Figure 47.1 shows some examples of data pairs and their correlation.

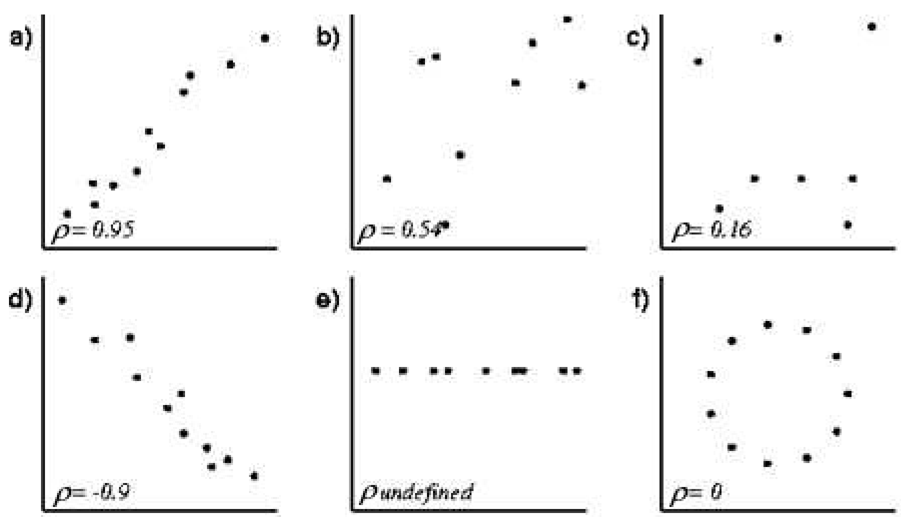


Figure 47.1

Practice Problems

Problem 47.1
If X and Y are independent and identically distributed with mean μ and variance σ^2, find $E[(X-Y)^2]$.

Problem 47.2
Two cards are drawn without replacement from a pack of cards. The random variable X measures the number of heart cards drawn, and the random variable Y measures the number of club cards drawn. Find the covariance and correlation of X and Y.

Problem 47.3
Suppose the joint pdf of X and Y is

$$f_{XY}(x,y) = \begin{cases} 1 & 0 < x < 1,\ x < y < x+1 \\ 0 & \text{otherwise} \end{cases}$$

Compute the covariance and correlation of X and Y..

Problem 47.4
Let X and Z be independent random variables with X uniformly distributed on $(-1,1)$ and Z uniformly distributed on $(0, 0.1)$. Let $Y = X^2 + Z$. Then X and Y are dependent.
(a) Find the joint pdf of X and Y.
(b) Find the covariance and the correlation of X and Y.

Problem 47.5
Let the random variable Θ be uniformly distributed on $[0, 2\pi]$. Consider the random variables $X = \cos\Theta$ and $Y = \sin\Theta$. Show that $Cov(X,Y) = 0$ even though X and Y are dependent. This means that there is a weak relationship between X and Y.

Problem 47.6
If X_1, X_2, X_3, X_4 are (pairwise) uncorrelated random variables each having mean 0 and variance 1, compute the correlations of
(a) $X_1 + X_2$ and $X_2 + X_3$.
(b) $X_1 + X_2$ and $X_3 + X_4$.

Problem 47.7
Let X be the number of 1's and Y the number of 2's that occur in n rolls of a fair die. Compute $\mathrm{Cov}(X, Y)$.

Problem 47.8
Let X be uniformly distributed on $[-1, 1]$ and $Y = X^2$. Show that X and Y are uncorrelated even though Y depends functionally on X (the strongest form of dependence).

Problem 47.9
Let X and Y be continuous random variables with joint pdf

$$f_{XY}(x,y) = \begin{cases} 3x & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $\mathrm{Cov}(X, Y)$ and $\rho(X, Y)$.

Problem 47.10
Suppose that X and Y are random variables with $Cov(X, Y) = 3$. Find $Cov(2X - 5, 4Y + 2)$.

Problem 47.11 ‡
An insurance policy pays a total medical benefit consisting of two parts for each claim. Let X represent the part of the benefit that is paid to the surgeon, and let Y represent the part that is paid to the hospital. The variance of X is 5000, the variance of Y is 10,000, and the variance of the total benefit, $X + Y$, is 17,000.
Due to increasing medical costs, the company that issues the policy decides to increase X by a flat amount of 100 per claim and to increase Y by 10% per claim.
Calculate the variance of the total benefit after these revisions have been made.

Problem 47.12 ‡
The profit for a new product is given by $Z = 3X - Y - 5$. X and Y are independent random variables with Var(X) = 1 and Var(Y) = 2.
What is the variance of Z?

Problem 47.13 ‡
A company has two electric generators. The time until failure for each generator follows an exponential distribution with mean 10. The company will begin using the second generator immediately after the first one fails.
What is the variance of the total time that the generators produce electricity?

Problem 47.14 ‡
A joint density function is given by

$$f_{XY}(x,y) = \begin{cases} kx & 0 < x, y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $\text{Cov}(X,Y)$

Problem 47.15 ‡
Let X and Y be continuous random variables with joint density function

$$f_{XY}(x,y) = \begin{cases} \frac{8}{3}xy & 0 \le x \le 1, x \le y \le 2x \\ 0 & \text{otherwise} \end{cases}$$

Find $\text{Cov}(X,Y)$

Problem 47.16 ‡
Let X and Y denote the values of two stocks at the end of a five-year period. X is uniformly distributed on the interval (0, 12) . Given $X = x$, Y is uniformly distributed on the interval $(0,x)$.
Determine $Cov(X,Y)$ according to this model.

Problem 47.17 ‡
Let X denote the size of a surgical claim and let Y denote the size of the associated hospital claim. An actuary is using a model in which $E(X) = 5, E(X^2) = 27.4, E(Y) = 7, E(Y^2) = 51.4$, and $Var(X+Y) = 8$.
Let $C_1 = X+Y$ denote the size of the combined claims before the application of a 20% surcharge on the hospital portion of the claim, and let C_2 denote the size of the combined claims after the application of that surcharge.
Calculate $Cov(C_1, C_2)$.

Problem 47.18 ‡
Claims filed under auto insurance policies follow a normal distribution with mean 19,400 and standard deviation 5,000.
What is the probability that the average of 25 randomly selected claims exceeds 20,000 ?

Problem 47.19
Let X and Y be two independent random variables with densities

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{2} & 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Write down the joint pdf $f_{XY}(x, y)$.
(b) Let $Z = X + Y$. Find the pdf $f_Z(a)$. Simplify as much as possible.
(c) Find the expectation $E(X)$ and variance $\text{Var}(X)$. Repeat for Y.
(d) Compute the expectation $E(Z)$ and the variance $\text{Var}(Z)$.

Problem 47.20
Let X and Y be two random variables with joint pdf

$$f_{XY}(x, y) = \begin{cases} \frac{1}{2} & x > 0, y > 0, x + y < 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Let $Z = X + Y$. Find the pdf of Z.
(b) Find the pdf of X and that of Y.
(c) Find the expectation and variance of X.
(d) Find the covariance $\text{Cov}(X, Y)$.

Problem 47.21
Let X and Y be discrete random variables with joint distribution defined by the following table

Y\ X	2	3	4	5	$p_Y(y)$
0	0.05	0.05	0.15	0.05	0.30
1	0.40	0	0	0	0.40
2	0.05	0.15	0.10	0	0.30
$p_X(x)$	0.50	0.20	0.25	0.05	1

For this joint distribution, $E(X) = 2.85,\ E(Y) = 1$. Calculate Cov(X, Y).

Problem 47.22
Let X and Y be two random variables with joint density function

$$f_{XY}(x, y) = \begin{cases} 1 & 0 < y < 1 - |x|, -1 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $\text{Var}(X)$.

Problem 47.23
Let X_1, X_2, X_3 be uniform random variables on the interval $(0,1)$ with $\text{Cov}(X_i, X_j) = \frac{1}{24}$ for $i, j \in \{1,2,3\},\ i \neq j$. Calculate the variance of $X_1 + 2X_2 - X_3$.

Problem 47.24
Let X and X be discrete random variables with joint probability function $p_{XY}(x,y)$ given by the following table:

X\ Y	0	1	$p_X(x)$
0	0	0.20	0.20
1	0.40	0.20	0.60
2	0.20	0	0.20
$p_Y(y)$	0.60	0.40	1

Find the variance of $Y - X$.

Problem 47.25
Let X and Y be two independent identically distributed normal random variables with mean 1 and variance 1. Find c so that $E[c|X - Y|] = 1$.

Problem 47.26
Let X, Y and Z be random variables with means 1,2 and 3, respectively, and variances 4,5, and 9, respectively. Also, $\text{Cov}(X,Y) = 2, \text{Cov}(X,Z) = 3$, and $\text{Cov}(Y,Z) = 1$. What are the mean and variance, respectively, of the random variable $W = 3X + 2Y - Z$?

Problem 47.27
Let X_1, X_2, and X_3 be independent random variables each with mean 0 and variance 1. Let $X = 2X_1 - X_3$ and $Y = 2X_2 + X_3$. Find $\rho(X,Y)$.

Problem 47.28
The coefficient of correlation between random variables X and Y is $\frac{1}{3}$, and $\sigma_X^2 = a, \sigma_Y^2 = 4a$. The random variable Z is defined to be $Z = 3X - 4Y$, and it is found that $\sigma_Z^2 = 114$. Find a.

Problem 47.29
Given n independent random variables $X_1, X_2, \cdots, X_n$ each having the same variance σ^2. Define $U = 2X_1 + X_2 + \cdots + X_{n-1}$ and $V = X_2 + X_3 + \cdots + X_{n-1} + 2X_n$. Find $\rho(U,V)$.

Problem 47.30
The following table gives the joint probability distribution of two random variables X and Y.

X \ Y	0	1	2
0	0.25	0.08	0.05
1	0.12	0.20	0.10
2	0.03	0.07	0.10

(a) Give the marginal distributions of X and Y.
(b) Find $E(X)$ and $E(Y)$.
(c) Find $\mathrm{Cov}(X, Y)$.
(d) Find $E(100X + 75Y)$.

48 Conditional Expectation

Since conditional probability measures are probabilitiy measures (that is, they possess all of the properties of unconditional probability measures), conditional expectations inherit all of the properties of regular expectations. Let X and Y be random variables. We define **conditional expectation** of X given that $Y = y$ by

$$\begin{aligned} E(X|Y=y\} =& \sum_x x\Pr(X=x|Y=y) \\ =& \sum_x xp_{X|Y}(x|y) \end{aligned}$$

where $p_{X|Y}$ is the conditional probability mass function of X, given that $Y = y$ which is given by

$$p_{X|Y}(x|y) = \Pr(X=x|Y=y) = \frac{\Pr(x,y)}{p_Y(y)}.$$

In the continuous case we have

$$E(X|Y=y) = \int_{-\infty}^{\infty} xf_{X|Y}(x|y)dx$$

where

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}.$$

Example 48.1
Suppose X and Y are discrete random variables with values 1, 2, 3, 4 and joint p.m.f. given by

$$f(x,y) = \begin{cases} \frac{1}{16} & \text{if } x = y \\ \frac{2}{16} & \text{if } x < y \\ 0 & \text{if } x > y \end{cases}$$

for $x, y = 1, 2, 3, 4$.
(a) Find the joint probability distribution of X and Y.
(b) Find the conditional expectation of Y given that $X = 3$.

Solution.
(a) The joint probability distribution is given in tabular form

X \ Y	1	2	3	4	$p_X(x)$
1	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{7}{16}$
2	0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{5}{16}$
3	0	0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$
4	0	0	0	$\frac{1}{16}$	$\frac{1}{16}$
$p_Y(y)$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{7}{16}$	1

(b) We have

$$\begin{aligned}
E(Y|X=3) =& \sum_{y=1}^{4} y p_{Y|X}(y|3) \\
=& \frac{p_{XY}(3,1)}{p_X(3)} + \frac{2p_{XY}(3,2)}{p_X(3)} + \frac{3p_{XY}(3,3)}{p_X(3)} + \frac{4p_{XY}(3,4)}{p_X(3)} \\
=& 3 \cdot \frac{1}{3} + 4 \cdot \frac{2}{3} = \frac{11}{3} \blacksquare
\end{aligned}$$

Example 48.2
Suppose that the joint density of X and Y is given by

$$f_{XY}(x,y) = \frac{e^{-\frac{x}{y}} e^{-y}}{y}, \quad x, y > 0.$$

Compute $E(X|Y=y)$.

Solution.
The conditional density is found as follows

$$\begin{aligned}
f_{X|Y}(x|y) =& \frac{f_{XY}(x,y)}{f_Y(y)} \\
=& \frac{f_{XY}(x,y)}{\int_{-\infty}^{\infty} f_{XY}(x,y)dx} \\
=& \frac{(1/y)e^{-\frac{x}{y}} e^{-y}}{\int_0^{\infty} (1/y)e^{-\frac{x}{y}} e^{-y} dx} \\
=& \frac{(1/y)e^{-\frac{x}{y}}}{\int_0^{\infty} (1/y)e^{-\frac{x}{y}} dx} \\
=& \frac{1}{y} e^{-\frac{x}{y}}
\end{aligned}$$

Hence,

$$E(X|Y=y) = \int_0^\infty \frac{x}{y}e^{-\frac{x}{y}}dx = -\left[xe^{-\frac{x}{y}}\Big|_0^\infty - \int_0^\infty e^{-\frac{x}{y}}dx\right]$$
$$= -\left[xe^{-\frac{x}{y}} + ye^{-\frac{x}{y}}\right]_0^\infty = y \blacksquare$$

Example 48.3
Let Y be a random variable with a density f_Y given by

$$f_Y(y) = \begin{cases} \frac{\alpha-1}{y^\alpha} & y > 1 \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha > 1$. Given $Y = y$, let X be a random variable which is Uniformly distributed on $(0, y)$.
(a) Find the marginal distribution of X.
(b) Calculate $E(Y|X = x)$ for every $x > 0$.

Solution.
The joint density function is given by

$$f_{XY}(x,y) = \begin{cases} \frac{\alpha-1}{y^{\alpha+1}} & 0 < x < y,\ y > 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Observe that X only takes positive values, thus $f_X(x) = 0,\ x \le 0$. For $0 < x < 1$ we have

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_1^\infty f_{XY}(x,y)dy = \frac{\alpha-1}{\alpha}.$$

For $x \ge 1$ we have

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_x^\infty f_{XY}(x,y)dy = \frac{\alpha-1}{\alpha x^\alpha}.$$

(b) For $0 < x < 1$ we have

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{\alpha}{y^{\alpha+1}}, \quad y > 1.$$

Hence,

$$E(Y|X=x) = \int_1^\infty \frac{y\alpha}{y^{\alpha+1}}dy = \alpha\int_1^\infty \frac{dy}{y^\alpha} = \frac{\alpha}{\alpha-1}.$$

If $x \geq 1$ then

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{\alpha x^\alpha}{y^{\alpha+1}}, \quad y > x.$$

Hence,

$$E(Y|X = x) = \int_x^\infty y \frac{\alpha x^\alpha}{y^{\alpha+1}} dy = \frac{\alpha x}{\alpha - 1} \blacksquare$$

Notice that if X and Y are independent then $p_{X|Y}(x|y) = \Pr(x)$ so that $E(X|Y = y) = E(X)$.
Now, for any function $g(x)$, the conditional expected value of g given $Y = y$ is, in the continuous case,

$$E(g(X)|Y = y) = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y) dx$$

if the integral exists. For the discrete case, we have a sum instead of an integral. That is, the conditional expectation of g given $Y = y$ is

$$E(g(X)|Y = y) = \sum_x g(x) p_{X|Y}(x|y).$$

The proof of this result is identical to the unconditional case.
Next, let $\phi_X(y) = E(X|Y = y)$ denote the function of the random variable Y whose value at $Y = y$ is $E(X|Y = y)$. Clearly, $\phi_X(y)$ is a random variable. We denote this random variable by $E(X|Y)$. The expectation of this random variable is just the expectation of X as shown in the following theorem.

Theorem 48.1 *(Double Expectation Property)*

$$E(X) = E(E(X|Y))$$

Proof.
We give a proof in the case X and Y are continuous random variables.

$$\begin{aligned}
E(E(X|Y)) &= \int_{-\infty}^{\infty} E(X|Y=y)f_Y(y)dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} xf_{X|Y}(x|y)dx\right) f_Y(y)dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf_{X|Y}(x|y)f_Y(y)dxdy \\
&= \int_{-\infty}^{\infty} x\int_{-\infty}^{\infty} f_{XY}(x,y)dydx \\
&= \int_{-\infty}^{\infty} xf_X(x)dx = E(X) \blacksquare
\end{aligned}$$

Computing Probabilities by Conditioning
Suppose we want to know the probability of some event, A. Suppose also that knowing Y gives us some useful information about whether or not A occurred.
Define an indicator random variable

$$X = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A \text{ does not occur} \end{cases}$$

Then

$$\Pr(A) = E(X)$$

and for any random variable Y

$$E(X|Y=y) = \Pr(A|Y=y).$$

Thus, by the double expectation property we have

$$\begin{aligned}
\Pr(A) =&E(X) = \sum_y E(X|Y=y)\Pr(Y=y) \\
=&\sum_y \Pr(A|Y=y)p_Y(y)
\end{aligned}$$

in the discrete case and

$$\Pr(A) = \int_{-\infty}^{\infty} \Pr(A|Y=y)f_Y(y)dy$$

in the continuous case.

The Conditional Variance

Next, we introduce the concept of conditional variance. Just as we have defined the conditional expectation of X given that $Y = y$, we can define the conditional variance of X given Y as follows

$$\text{Var}(X|Y = y) = E[(X - E(X|Y))^2|Y = y].$$

Note that the conditional variance is a random variable since it is a function of Y.

Proposition 48.1

Let X and Y be random variables. Then
(a) $\text{Var}(X|Y) = E(X^2|Y) - [E(X|Y)]^2$.
(b) $E(\text{Var}(X|Y)) = E[E(X^2|Y) - (E(X|Y))^2] = E(X^2) - E[(E(X|Y))^2]$.
(c) $\text{Var}(E(X|Y)) = E[(E(X|Y))^2] - (E(X))^2$.
(d) Law of Total Variance: $\text{Var}(X) = E[\text{Var}(X|Y)] + \text{Var}(E(X|Y))$.

Proof.

(a) We have

$$\begin{aligned}\text{Var}(X|Y) =& E[(X - E(X|Y))^2|Y]\\ =& E[(X^2 - 2XE(X|Y) + (E(X|Y))^2|Y]\\ =& E(X^2|Y) - 2E(X|Y)E(X|Y) + (E(X|Y))^2\\ =& E(X^2|Y) - [E(X|Y)]^2.\end{aligned}$$

(b) Taking E of both sides of the result in (a) we find

$$E(\text{Var}(X|Y)) = E[E(X^2|Y) - (E(X|Y))^2] = E(X^2) - E[(E(X|Y))^2].$$

(c) Since $E(E(X|Y)) = E(X)$ we have

$$\text{Var}(E(X|Y)) = E[(E(X|Y))^2] - (E(X))^2.$$

(d) The result follows by adding the two equations in (b) and (c) ■

Conditional Expectation and Prediction

One of the most important uses of conditional expectation is in estimation theory. Let us begin this discussion by asking: What constitutes a good estimator? An obvious answer is that the estimate be close to the true value. Suppose that we are in a situation where the value of a random variable is observed and then, based on the observed, an attempt is made to predict the value of a second random variable Y. Let $g(X)$ denote the predictor, that is, if X is observed to be equal to x, then $g(x)$ is our prediction for the value of Y. So the question is of choosing g in such a way $g(X)$ is close to Y. One possible criterion for closeness is to choose g so as to minimize $E[(Y-g(X))^2]$. Such a minimizer will be called **minimum mean square estimate** (MMSE) of Y given X. The following theorem shows that the MMSE of Y given X is just the conditional expectation $E(Y|X)$.

Theorem 48.2

$$\min_g E[(Y-g(X))^2] = E(Y-E(Y|X)).$$

Proof.
We have

$$\begin{aligned}E[(Y-g(X))^2] =& E[(Y-E(Y|X)+E(Y|X)-g(X))^2]\\ =& E[(Y-E(Y|X))^2]+E[(E(Y|X)-g(X))^2]\\ &+2E[(Y-E(Y|X))(E(Y|X)-g(X))].\end{aligned}$$

Using the fact that the expression $h(X)=E(Y|X)-g(X)$ is a function of X and thus can be treated as a constant we have

$$\begin{aligned}E[(Y-E(Y|X))h(X)] =& E[E[(Y-E(Y|X))h(X)|X]]\\ =& E[(h(X)E[Y-E(Y|X)|X]]\\ =& E[h(X)[E(Y|X)-E(Y|X)]]=0\end{aligned}$$

for all functions g. Thus,

$$E[(Y-g(X))^2]=E[(Y-E(Y|X))^2]+E[(E(Y|X)-g(X))^2].$$

The first term on the right of the previous equation is not a function of g. Thus, the right hand side expression is minimized when $g(X)=E(Y|X)$ ■

Example 48.4
Let X and Y be random variables with joint pdf

$$f_{XY} = \begin{cases} \frac{\alpha-1}{y^\alpha} & y > 1 \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha > 1$. Determine the best estimator $g(x)$ of Y given X.

Solution.
Given $Y = y$, let X be a random variable which is Uniformly distributed on $(0, y)$. From Example 48.3, the best estimator is given by

$$g(x) = E(Y|X) = \begin{cases} \frac{\alpha}{\alpha-1}, & \text{if } x < 1 \\ \frac{\alpha x}{\alpha-1}, & \text{if } x \geq 1 \end{cases} \blacksquare$$

Practice Problems

Problem 48.1
Suppose that X and Y have joint distribution

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

Find $E(X|Y)$ and $E(Y|X)$.

Problem 48.2
Suppose that X and Y have joint distribution

$$f_{XY}(x,y) = \begin{cases} \frac{3y^2}{x^3} & 0 < y < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find $E(X),\ E(X^2), Var(X), E(Y|X), Var(Y|X), E[Var(Y|X)], Var[E(Y|X)],$ and $Var(Y)$.

Problem 48.3
Let X and Y be independent exponentially distributed random variables with parameters μ and λ respectively. Using conditioning, find $\Pr(X > Y)$.

Problem 48.4
Let X be uniformly distributed on $[0,1]$. Find $E(X|X > 0.5)$.

Problem 48.5
Let X and Y be discrete random variables with conditional density function

$$f_{Y|X}(y|2) = \begin{cases} 0.2 & y = 1 \\ 0.3 & y = 2 \\ 0.5 & y = 3 \\ 0 & \text{otherwise.} \end{cases}$$

Compute $E(Y|X = 2)$.

Problem 48.6
Suppose that X and Y have joint distribution

$$f_{XY}(x,y) = \begin{cases} \frac{21}{4}x^2y & x^2 < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find $E(Y|X)$.

Problem 48.7
Suppose that X and Y have joint distribution

$$f_{XY}(x,y) = \begin{cases} \frac{21}{4}x^2y & x^2 < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find $E(Y)$ in two ways.

Problem 48.8
Suppose that $E(X|Y) = 18 - \frac{3}{5}Y$ and $E(Y|X) = 10 - \frac{1}{3}X$. Find $E(X)$ and $E(Y)$.

Problem 48.9
Let X be an exponential random variable with $\lambda = 5$ and Y a uniformly distributed random variable on $(-3, X)$. Find $E(Y)$.

Problem 48.10
In a mall, a survey found that the number of people who pass by JCPenney between 4:00 and 5:00 pm is a Poisson random variable with parameter $\lambda = 100$. Assume that each person may enter the store, independently of the other person, with a given probability $p = 0.15$. What is the expected number of people who enter the store during the given period?

Problem 48.11
Let X and Y be discrete random variables with joint probability mass function defined by the following table

$X\backslash Y$	1	2	3	$p_X(x)$
1	1/9	1/9	0	2/9
2	1/3	0	1/6	1/2
3	1/9	1/18	1/9	5/18
$p_Y(y)$	5/9	1/6	5/18	1

Compute $E(X|Y = i)$ for $i = 1, 2, 3$. Are X and Y independent?

Problem 48.12 ‡
A diagnostic test for the presence of a disease has two possible outcomes: 1 for disease present and 0 for disease not present. Let X denote the disease

state of a patient, and let Y denote the outcome of the diagnostic test. The joint probability function of X and Y is given by:

$$\begin{aligned}\Pr(X=0,Y=0)=&0.800\\ \Pr(X=1,Y=0)=&0.050\\ \Pr(X=0,Y=1)=&0.025\\ \Pr(X=1,Y=1)=&0.125.\end{aligned}$$

Calculate $Var(Y|X=1)$.

Problem 48.13 ‡
The stock prices of two companies at the end of any given year are modeled with random variables X and Y that follow a distribution with joint density function

$$f_{XY}(x,y)=\begin{cases}2x & 0<x<1, x<y<x+1\\ 0 & \text{otherwise.}\end{cases}$$

What is the conditional variance of Y given that $X=x$?

Problem 48.14 ‡
An actuary determines that the annual numbers of tornadoes in counties P and Q are jointly distributed as follows:

$X\backslash Y$	0	1	2	$P_X(x)$
0	0.12	0.13	0.05	0.30
1	0.06	0.15	0.15	0.36
2	0.05	0.12	0.10	0.27
3	0.02	0.03	0.02	0.07
$p_Y(y)$	0.25	0.43	0.32	1

where X is the number of tornadoes in county Q and Y that of county P. Calculate the conditional variance of the annual number of tornadoes in county Q, given that there are no tornadoes in county P.

Problem 48.15
Let X be a random variable with mean 3 and variance 2, and let Y be a random variable such that for every x, the conditional distribution of Y given $X=x$ has a mean of x and a variance of x^2. What is the variance of the marginal distribution of Y?

Problem 48.16
Let X and Y be two continuous random variables with joint density function

$$f_{XY}(x,y) = \begin{cases} 2 & 0 < x < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

For $0 < x < 1$, find $\text{Var}(Y|X = x)$.

Problem 48.17
Suppose that the number of stops X in a day for a UPS delivery truck driver is Poisson with mean λ and that the expected distance driven by the driver Y, given that there are $X = x$ stops, has a normal distribution with a mean of αx miles, and a standard deviation of βx miles. Find the mean and variance of the number of miles driven per day.

Problem 48.18 ‡
The joint probability density for X and Y is

$$f(x,y) = \begin{cases} 2e^{-(x+2y)}, & \text{for } x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the variance of Y given that $X > 3$ and $Y > 3$.

Problem 48.19 ‡
The number of workplace injuries, N, occurring in a factory on any given day is Poisson distributed with mean λ. The parameter λ is a random variable that is determined by the level of activity in the factory, and is uniformly distributed on the interval $[0, 3]$.
Calculate $\text{Var}(N)$.

Problem 48.20 ‡
A fair die is rolled repeatedly. Let X be the number of rolls needed to obtain a 5 and Y the number of rolls needed to obtain a 6. Calculate $E(X|Y = 2)$.

Problem 48.21 ‡
A driver and a passenger are in a car accident. Each of them independently has probability 0.3 of being hospitalized. When a hospitalization occurs, the loss is uniformly distributed on [0, 1]. When two hospitalizations occur, the losses are independent.
Calculate the expected number of people in the car who are hospitalized, given that the total loss due to hospitalizations from the accident is less than 1.

Problem 48.22 ‡
New dental and medical plan options will be offered to state employees next year. An actuary uses the following density function to model the joint distribution of the proportion X of state employees who will choose Dental Option 1 and the proportion Y who will choose Medical Option 1 under the new plan options:

$$f(x,y) = \begin{cases} 0.50 & \text{for } 0 < x, y < 0.5 \\ 1.25 & \text{for } 0 < x < 0.5, 0.5 < y < 1 \\ 1.50 & \text{for } 0.5 < x < 1, 0 < y < 0.5 \\ 0.75 & \text{for } 0.5 < x < 1, 0.5 < y < 1. \end{cases}$$

Calculate $\text{Var}(Y|X = 0.75)$.

Problem 48.23 ‡
A motorist makes three driving errors, each independently resulting in an accident with probability 0.25.
Each accident results in a loss that is exponentially distributed with mean 0.80. Losses are mutually independent and independent of the number of accidents. The motorist's insurer reimburses 70% of each loss due to an accident.
Calculate the variance of the total unreimbursed loss the motorist experiences due to accidents resulting from these driving errors.

Problem 48.24 ‡
The number of hurricanes that will hit a certain house in the next ten years is Poisson distributed with mean 4.
Each hurricane results in a loss that is exponentially distributed with mean 1000. Losses are mutually independent and independent of the number of hurricanes.
Calculate the variance of the total loss due to hurricanes hitting this house in the next ten years.

Problem 48.25
Let X and Y be random variables with joint pdf

$$f_{XY} = \begin{cases} \alpha xy & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha > 0$. Determine the best estimator $g(x)$ of Y given X.

49 Moment Generating Functions

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

$$M_X(t) = E[e^{tX}]$$

provided that the expectation exists for t in some neighborhood of 0.
For a discrete random variable with a pmf $\Pr(x)$ we have

$$M_X(t) = \sum_x e^{tx}\Pr(x)$$

and for a continuous random variable with pdf f,

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx.$$

Example 49.1
Let X be a discrete random variable with pmf given by the following table

x	1	2	3	4	5
Pr(x)	0.15	0.20	0.40	0.15	0.10

Find $M_X(t)$.

Solution.
We have

$$M_X(t) = 0.15e^t + 0.20e^{2t} + 0.40e^{3t} + 0.15e^{4t} + 0.10e^{5t} \blacksquare$$

Example 49.2
Let X be the uniform random variable on the interval $[a, b]$. Find $M_X(t)$.

Solution.
We have

$$M_X(t) = \int_a^b \frac{e^{tx}}{b-a}dx = \frac{1}{t(b-a)}[e^{tb} - e^{ta}] \blacksquare$$

As the name suggests, the moment generating function can be used to generate moments $E(X^n)$ for $n = 1, 2, \cdots$. Our first result shows how to use the moment generating function to calculate moments.

Proposition 49.1

$$E(X^n) = M_X^n(0)$$

where

$$M_X^n(0) = \left.\frac{d^n}{dt^n} M_X(t)\right|_{t=0}.$$

Proof.
We prove the result for a continuous random variable X with pdf f. The discrete case is shown similarly. In what follows we always assume that we can differentiate under the integral sign. This interchangeability of differentiation and expectation is not very limiting, since all of the distributions we will consider enjoy this property. We have

$$\begin{aligned}\frac{d}{dt}M_X(t) =&\frac{d}{dt}\int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_{-\infty}^{\infty}\left(\frac{d}{dt}e^{tx}\right) f(x)dx\\ =&\int_{-\infty}^{\infty} xe^{tx}f(x)dx = E[Xe^{tX}]\end{aligned}$$

Hence,

$$\frac{d}{dt}M_X(t)\mid_{t=0} = E[Xe^{tX}]\mid_{t=0} = E(X).$$

By induction on n we find

$$\frac{d^n}{dt^n}M_X(t)\mid_{t=0} = E[X^n e^{tX}]\mid_{t=0} = E(X^n) \blacksquare$$

We next compute $M_X(t)$ for some common distributions.

Example 49.3
Let X be a binomial random variable with parameters n and p. Find the expected value and the variance of X using moment generating functions.

Solution.
We can write

$$\begin{aligned}M_X(t) =&E(e^{tX}) = \sum_{k=0}^{n} e^{tk}{}_nC_k p^k(1-p)^{n-k}\\ =&\sum_{k=0}^{n} {}_nC_k(pe^t)^k(1-p)^{n-k} = (pe^t+1-p)^n\end{aligned}$$

Differentiating yields

$$\frac{d}{dt}M_X(t) = npe^t(pe^t + 1 - p)^{n-1}$$

Thus

$$E(X) = \frac{d}{dt}M_X(t) \mid_{t=0} = np.$$

To find $E(X^2)$, we differentiate a second time to obtain

$$\frac{d^2}{dt^2}M_X(t) = n(n-1)p^2e^{2t}(pe^t + 1 - p)^{n-2} + npe^t(pe^t + 1 - p)^{n-1}.$$

Evaluating at $t = 0$ we find

$$E(X^2) = M_X''(0) = n(n-1)p^2 + np.$$

Observe that this implies the variance of X is

$$Var(X) = E(X^2) - (E(X))^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p) \blacksquare$$

Example 49.4
Let X be a Poisson random variable with parameter λ. Find the expected value and the variance of X using moment generating functions.

Solution.
We can write

$$\begin{aligned} M_X(t) =& E(e^{tX}) = \sum_{n=0}^{\infty} \frac{e^{tn}e^{-\lambda}\lambda^n}{n!} = e^{-\lambda}\sum_{n=0}^{\infty}\frac{e^{tn}\lambda^n}{n!} \\ =& e^{-\lambda}\sum_{n=0}^{\infty}\frac{(\lambda e^t)^n}{n!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda(e^t-1)} \end{aligned}$$

Differentiating for the first time we find

$$M_X'(t) = \lambda e^t e^{\lambda(e^t-1)}.$$

Thus,

$$E(X) = M_X'(0) = \lambda.$$

Differentiating a second time we find

$$M''_X(t) = (\lambda e^t)^2 e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)}.$$

Hence,

$$E(X^2) = M''_X(0) = \lambda^2 + \lambda.$$

The variance is then

$$Var(X) = E(X^2) - (E(X))^2 = \lambda \blacksquare$$

Example 49.5
Let X be an exponential random variable with parameter λ. Find the expected value and the variance of X using moment generating functions.

Solution.
We can write

$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx = \lambda\int_0^\infty e^{-(\lambda-t)x}dx = \frac{\lambda}{\lambda-t}$$

where $t < \lambda$. Differentiation twice yields

$$M'_X(t) = \frac{\lambda}{(\lambda-t)^2} \quad \text{and} \quad M''_X(t) = \frac{2\lambda}{(\lambda-t)^3}.$$

Hence,

$$E(X) = M'_X(0) = \frac{1}{\lambda} \quad \text{and} \quad E(X^2) = M''_X(0) = \frac{2}{\lambda^2}.$$

The variance of X is given by

$$Var(X) = E(X^2) - (E(X))^2 = \frac{1}{\lambda^2} \blacksquare$$

Moment generating functions are also useful in establishing the distribution of sums of independent random variables. To see this, the following two observations are useful. Let X be a random variable, and let a and b be finite constants. Then,

$$\begin{aligned} M_{aX+b}(t) =& E[e^{t(aX+b)}] = E[e^{bt}e^{(at)X}] \\ =& e^{bt}E[e^{(at)X}] = e^{bt}M_X(at) \end{aligned}$$

Example 49.6
Let X be a normal random variable with parameters μ and σ^2. Find the expected value and the variance of X using moment generating functions.

Solution.
First we find the moment of a standard normal random variable with parameters 0 and 1. We can write

$$\begin{aligned}
M_Z(t) =& E(e^{tZ}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tz}e^{-\frac{z^2}{2}}dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} exp\left\{-\frac{(z^2-2tz)}{2}\right\}dz\\
=&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} exp\left\{-\frac{(z-t)^2}{2}+\frac{t^2}{2}\right\}dz = e^{\frac{t^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(z-t)^2}{2}}dz = e^{\frac{t^2}{2}}
\end{aligned}$$

Now, since $X = \mu + \sigma Z$ we have

$$\begin{aligned}
M_X(t) =& E(e^{tX}) = E(e^{t\mu+t\sigma Z}) = E(e^{t\mu}e^{t\sigma Z}) = e^{t\mu}E(e^{t\sigma Z})\\
=& e^{t\mu}M_Z(t\sigma) = e^{t\mu}e^{\frac{\sigma^2t^2}{2}} = exp\left\{\frac{\sigma^2t^2}{2}+\mu t\right\}
\end{aligned}$$

By differentiation we obtain

$$M_X'(t) = (\mu + t\sigma^2)exp\left\{\frac{\sigma^2t^2}{2}+\mu t\right\}$$

and

$$M_X''(t) = (\mu + t\sigma^2)^2exp\left\{\frac{\sigma^2t^2}{2}+\mu t\right\} + \sigma^2exp\left\{\frac{\sigma^2t^2}{2}+\mu t\right\}$$

and thus

$$E(X) = M_X'(0) = \mu \text{ and } E(X^2) = M_X''(0) = \mu^2 + \sigma^2$$

The variance of X is

$$Var(X) = E(X^2) - (E(X))^2 = \sigma^2 \blacksquare$$

Next, suppose $X_1, X_2, \cdots, X_N$ are independent random variables. Then, the moment generating function of $Y = X_1 + \cdots + X_N$ is

$$\begin{aligned}
M_Y(t) =& E(e^{t(X_1+X_2+\cdots+X_n)}) = E(e^{X_1t}\cdots e^{X_Nt})\\
=& \prod_{k=1}^{N} E(e^{X_kt}) = \prod_{k=1}^{N} M_{X_k}(t)
\end{aligned}$$

where the next-to-last equality follows from Proposition 46.5.

Another important property is that the moment generating function uniquely determines the distribution. That is, if random variables X and Y both have moment generating functions $M_X(t)$ and $M_Y(t)$ that exist in some neighborhood of zero and if $M_X(t) = M_Y(t)$ for all t in this neighborhood, then X and Y have the same distributions.
The general proof of this is an inversion problem involving Laplace transform theory and is omitted. However, We will prove the claim here in a simplified setting.
Suppose X and Y are two random variables with common range $\{0, 1, 2, \cdots, n\}$. Moreover, suppose that both variables have the same moment generating function. That is,

$$\sum_{x=0}^{n} e^{tx} p_X(x) = \sum_{y=0}^{n} e^{ty} p_Y(y).$$

For simplicity, let $s = e^t$ and $c_i = p_X(i) - p_Y(i)$ for $i = 0, 1, \cdots, n$. Then

$$\begin{aligned}
0 &= \sum_{x=0}^{n} e^{tx} p_X(x) - \sum_{y=0}^{n} e^{ty} p_Y(y) \\
0 &= \sum_{x=0}^{n} s^x p_X(x) - \sum_{y=0}^{n} s^y p_Y(y) \\
0 &= \sum_{i=0}^{n} s^i p_X(i) - \sum_{i=0}^{n} s^i p_Y(i) \\
0 &= \sum_{i=0}^{n} s^i [p_X(i) - p_Y(i)] \\
0 &= \sum_{i=0}^{n} c_i s^i, \quad \forall s > 0.
\end{aligned}$$

The above is simply a polynomial in s with coefficients $c_0, c_1, \cdots, c_n$. The only way it can be zero for all values of s is if $c_0 = c_1 = \cdots = c_n = 0$.
That is

$$p_X(i) = p_Y(i), \quad i = 0, 1, 2, \cdots, n.$$

So probability mass functions for X and Y are exactly the same.

Example 49.7
If X and Y are independent binomial random variables with parameters (n,p) and (m,p), respectively, what is the pmf of $X+Y$?

Solution.
We have

$$\begin{aligned} M_{X+Y}(t) =& M_X(t)M_Y(t)\\ =&(pe^t+1-p)^n(pe^t+1-p)^m\\ =&(pe^t+1-p)^{n+m}. \end{aligned}$$

Since $(pe^t+1-p)^{n+m}$ is the moment generating function of a binomial random variable having parameters $m+n$ and p, $X+Y$ is a binomial random variable with this same pmf ■

Example 49.8
If X and Y are independent Poisson random variables with parameters λ_1 and λ_2, respectively, what is the pmf of $X+Y$?

Solution.
We have

$$\begin{aligned} M_{X+Y}(t) =& M_X(t)M_Y(t)\\ =&e^{\lambda_1(e^t-1)}e^{\lambda_2(e^t-1)}\\ =&e^{(\lambda_1+\lambda_2)(e^t-1)}. \end{aligned}$$

Since $e^{(\lambda_1+\lambda_2)(e^t-1)}$ is the moment generating function of a Poisson random variable having parameter $\lambda_1+\lambda_2$, $X+Y$ is a Poisson random variable with this same pmf ■

Example 49.9
If X and Y are independent normal random variables with parameters (μ_1,σ_1^2) and (μ_2,σ_2^2), respectively, what is the distribution of $X+Y$?

Solution.
We have

$$\begin{aligned} M_{X+Y}(t) =& M_X(t)M_Y(t)\\ =&exp\left\{\frac{\sigma_1^2t^2}{2}+\mu_1 t\right\}\cdot exp\left\{\frac{\sigma_2^2t^2}{2}+\mu_2 t\right\}\\ =&exp\left\{\frac{(\sigma_1^2+\sigma_2^2)t^2}{2}+(\mu_1+\mu_2)t\right\} \end{aligned}$$

which is the moment generating function of a normal random variable with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$. Because the moment generating function uniquely determines the distribution then $X + Y$ is a normal random variable with the same distribution ■

Example 49.10 ‡
An insurance company insures two types of cars, economy cars and luxury cars. The damage claim resulting from an accident involving an economy car has normal $N(7, 1)$ distribution, the claim from a luxury car accident has normal $N(20, 6)$ distribution.
Suppose the company receives three claims from economy car accidents and one claim from a luxury car accident. Assuming that these four claims are mutually independent, what is the probability that the total claim amount from the three economy car accidents exceeds the claim amount from the luxury car accident?

Solution.
Let X_1, X_2, X_3 denote the claim amounts from the three economy cars, and X_4 the claim from the luxury car. Then we need to compute $\Pr(X_1 + X_2 + X_3 > X_4)$, which is the same as $\Pr(X_1 + X_2 + X_3 - X_4 > 0)$. Now, since the X_is are independent and normal with distribution $N(7, 1)$ (for $i = 1, 2, 3$) and $N(20, 6)$ for $i = 4$, the linear combination $X = X_1 + X_2 + X_3 - X_4$ has normal distribution with parameters $\mu = 7+7+7-20 = 1$ and $\sigma^2 = 1+1+1+6 = 9$. Thus, the probability we want is

$$\begin{aligned}
\Pr(X > 0) =& P\left(\frac{X-1}{\sqrt{9}} > \frac{0-1}{\sqrt{9}}\right) \\
=& \Pr(Z > -0.33) = 1 - \Pr(Z \leq -0.33) \\
=& \Pr(Z \leq 0.33) \approx 0.6293 \blacksquare
\end{aligned}$$

Joint Moment Generating Functions
For any random variables $X_1, X_2, \cdots, X_n$, the joint moment generating function is defined by

$$M(t_1, t_2, \cdots, t_n) = E(e^{t_1 X_1 + t_2 X_2 + \cdots + t_n X_n}).$$

Example 49.11
Let X and Y be two independent normal random variables with parameters

(μ_1, σ_1^2) and (μ_2, σ_2^2) respectively. Find the joint moment generating function of $X+Y$ and $X-Y$.

Solution.
The joint moment generating function is

$$\begin{aligned}M(t_1,t_2) =&E(e^{t_1(X+Y)+t_2(X-Y)}) = E(e^{(t_1+t_2)X+(t_1-t_2)Y})\\=&E(e^{(t_1+t_2)X})E(e^{(t_1-t_2)Y}) = M_X(t_1+t_2)M_Y(t_1-t_2)\\=&e^{(t_1+t_2)\mu_1+\frac{1}{2}(t_1+t_2)^2\sigma_1^2}e^{(t_1-t_2)\mu_2+\frac{1}{2}(t_1-t_2)^2\sigma_2^2}\\=&e^{(t_1+t_2)\mu_1+(t_1-t_2)\mu_2+\frac{1}{2}(t_1^2+t_2^2)\sigma_1^2+\frac{1}{2}(t_1^2+t_2^2)\sigma_2^2+t_1t_2(\sigma_1^2-\sigma_2^2)} \blacksquare\end{aligned}$$

Example 49.12
Let X and Y be two random variables with joint distribution function

$$f_{XY}(x,y) = \begin{cases} e^{-x-y} & x>0, y>0\\ 0 & \text{otherwise}\end{cases}$$

Find $E(XY), E(X), E(Y)$ and $\text{Cov}(X,Y)$.

Solution.
We note first that $f_{XY}(x,y) = f_X(x)f_Y(y)$ so that X and Y are independent. Thus, the moment generating function is given by

$$M(t_1,t_2) = E(e^{t_1X+t_2Y}) = E(e^{t_1X})E(e^{t_2Y}) = \frac{1}{1-t_1}\frac{1}{1-t_2}.$$

Thus,

$$E(XY) = \left.\frac{\partial^2}{\partial t_2\partial t_1}M(t_1,t_2)\right|_{(0,0)} = \left.\frac{1}{(1-t_1)^2(1-t_2)^2}\right|_{(0,0)} = 1$$

$$E(X) = \left.\frac{\partial}{\partial t_1}M(t_1,t_2)\right|_{(0,0)} = \left.\frac{1}{(1-t_1)^2(1-t_2)}\right|_{(0,0)} = 1$$

$$E(Y) = \left.\frac{\partial}{\partial t_2}M(t_1,t_2)\right|_{(0,0)} = \left.\frac{1}{(1-t_1)(1-t_2)^2}\right|_{(0,0)} = 1$$

and

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0 \blacksquare$$

Practice Problems

Problem 49.1
Let X be a discrete random variable with range $\{1, 2, \cdots, n\}$ so that its pmf is given by $p_X(j) = \frac{1}{n}$ for $1 \leq j \leq n$. Find $E(X)$ and $Var(X)$ using moment generating functions.

Problem 49.2
Let X be a geometric distribution function with $p_X(n) = p(1-p)^{n-1}$. Find the expected value and the variance of X using moment generating functions.

Problem 49.3
The following problem exhibits a random variable with no moment generating function. Let X be a random variable with pmf given by

$$p_X(n) = \frac{6}{\pi^2 n^2}, \quad n = 1, 2, 3, \cdots.$$

Show that $M_X(t)$ does not exist in any neighborhood of 0.

Problem 49.4
Let X be a gamma random variable with parameters α and λ. Find the expected value and the variance of X using moment generating functions.

Problem 49.5
Show that the sum of n independently exponential random variable each with paramter λ is a gamma random variable with parameters n and λ.

Problem 49.6
Let X be a random variable with pdf given by

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

Find $M_X(t)$.

Problem 49.7
Let X be an exponential random variable with paramter λ. Find the moment generating function of $Y = 3X - 2$.

Problem 49.8
Identify the random variable whose moment generating function is given by

$$M_X(t) = \left(\frac{3}{4}e^t + \frac{1}{4}\right)^{15}.$$

Problem 49.9
Identify the random variable whose moment generating function is given by

$$M_Y(t) = e^{-2t}\left(\frac{3}{4}e^{3t} + \frac{1}{4}\right)^{15}.$$

Problem 49.10 ‡
X and Y are independent random variables with common moment generating function $M(t) = e^{\frac{t^2}{2}}$. Let $W = X + Y$ and $Z = X - Y$. Determine the joint moment generating function, $M(t_1, t_2)$ of W and Z.

Problem 49.11 ‡
An actuary determines that the claim size for a certain class of accidents is a random variable, X, with moment generating function

$$M_X(t) = \frac{1}{(1 - 2500t)^4}$$

Determine the standard deviation of the claim size for this class of accidents.

Problem 49.12 ‡
A company insures homes in three cities, J, K, and L . Since sufficient distance separates the cities, it is reasonable to assume that the losses occurring in these cities are independent.
The moment generating functions for the loss distributions of the cities are:

$$\begin{aligned}
M_J(t) =&(1 - 2t)^{-3}\\
M_K(t) =&(1 - 2t)^{-2.5}\\
M_L(t) =&(1 - 2t)^{-4.5}
\end{aligned}$$

Let X represent the combined losses from the three cities. Calculate $E(X^3)$.

Problem 49.13 ‡
Let X_1, X_2, X_3 be independent discrete random variables with common probability mass function

$$\Pr(x) = \begin{cases} \frac{1}{3} & x = 0 \\ \frac{2}{3} & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Determine the moment generating function $M(t)$, of $Y = X_1X_2X_3$.

Problem 49.14 ‡
Two instruments are used to measure the height, h, of a tower. The error made by the less accurate instrument is normally distributed with mean 0 and standard deviation $0.0056h$. The error made by the more accurate instrument is normally distributed with mean 0 and standard deviation $0.0044h$.
Assuming the two measurements are independent random variables, what is the probability that their average value is within $0.005h$ of the height of the tower?

Problem 49.15
Let $X_1, X_2, \cdots, X_n$ be independent geometric random variables each with parameter p. Define $Y = X_1 + X_2 + \cdots X_n$.
(a) Find the moment generating function of $X_i,\ 1 \le i \le n$.
(b) Find the moment generating function of a negative binomial random variable with parameters (n, p).
(c) Show that Y defined above is a negative binomial random variable with parameters (n, p).

Problem 49.16
Let X be normally distributed with mean 500 and standard deviation 60 and Y be normally distributed with mean 450 and standard deviation 80. Suppose that X and Y are independent. Find $\Pr(X > Y)$.

Problem 49.17
Suppose a random variable X has moment generating function

$$M_X(t) = \left(\frac{2 + e^t}{3}\right)^9$$

Find the variance of X.

Problem 49.18
Let X be a random variable with density function

$$f(x) = \begin{cases} (k+1)x^2 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the moment generating function of X

Problem 49.19
If the moment generating function for the random variable X is $M_X(t) = \frac{1}{t+1}$, find $E[(X-2)^3]$.

Problem 49.20
Suppose that X is a random variable with moment generating function $M_X(t) = \sum_{j=0}^{\infty} \frac{e^{(tj-1)}}{j!}$. Find $\Pr(X = 2)$.

Problem 49.21
If X has a standard normal distribution and $Y = e^X$, what is the k-th moment of Y?

Problem 49.22
The random variable X has an exponential distribution with parameter b. It is found that $M_X(-b^2) = 0.2$. Find b.

Problem 49.23
Let X_1 and X_2 be two random variables with joint density function

$$f_{X_1X_1}(x_1, x_2) = \begin{cases} 1 & 0 < x_1 < 1, 0 < x_2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the moment generating function $M(t_1, t_2)$.

Problem 49.24
The moment generating function for the joint distribution of random variables X and Y is $M(t_1, t_2) = \frac{1}{3(1-t_2)} + \frac{2}{3}e^{t_1} \cdot \frac{2}{(2-t_2)}, \quad t_2 < 1$. Find $\mathrm{Var}(X)$.

Problem 49.25
Let X and Y be two independent random variables with moment generating functions

$$M_X(t) = e^{t^2+2t} \text{ and } M_Y(t) = e^{3t^2+t}$$

Determine the moment generating function of $X + 2Y$.

Problem 49.26
Let X_1 and X_2 be random variables with joint moment generating function

$$M(t_1, t_2) = 0.3 + 0.1e^{t_1} + 0.2e^{t_2} + 0.4e^{t_1+t_2}$$

What is $E(2X_1 - X_2)$?

Problem 49.27
Suppose X and Y are random variables whose joint distribution has moment generating function

$$M_{XY}(t_1, t_2) = \left(\frac{1}{4}e^{t_1} + \frac{3}{8}e^{t_2} + \frac{3}{8}\right)^{10}$$

for all t_1, t_2. Find the covariance between X and Y.

Problem 49.28
Independent random variables X, Y and Z are identically distributed. Let $W = X+Y$. The moment generating function of W is $M_W(t) = (0.7+0.3e^t)^6$. Find the moment generating function of $V = X + Y + Z$.

Problem 49.29 ‡
The value of a piece of factory equipment after three years of use is $100(0.5)^X$ where X is a random variable having moment generating function

$$M_X(t) = \tfrac{1}{1-2t} \text{ for } t < \tfrac{1}{2}.$$

Calculate the expected value of this piece of equipment after three years of use.

Problem 49.30 ‡
Let X and Y be identically distributed independent random variables such that the moment generating function of $X + Y$ is

$$M(t) = 0.09e^{-2t} + 0.24e^{-t} + 0.34 + 0.24e^t + 0.09e^{2t}, \quad -\infty < t < \infty.$$

Calculate $\Pr(X \leq 0)$.

Limit Theorems

Limit theorems are considered among the important results in probability theory. In this chapter, we consider two types of limit theorems. The first type is known as the **law of large numbers.** The law of large numbers describes how the average of a randomly selected sample from a large population is likely to be close to the average of the whole population.
The second type of limit theorems that we study is known as **central limit theorems**. Central limit theorems are concerned with determining conditions under which the sum of a large number of random variables has a probability distribution that is approximately normal.

50 The Law of Large Numbers

There are two versions of the law of large numbers: the weak law of large numbers and the strong law of numbers.

50.1 The Weak Law of Large Numbers

The law of large numbers is one of the fundamental theorems of statistics. One version of this theorem, the weak law of large numbers, can be proven in a fairly straightforward manner using Chebyshev's inequality, which is, in turn, a special case of the Markov inequality.
Our first result is known as **Markov's inequality**.

Proposition 50.1 *(Markov's Inequality)*
If $X \geq 0$ and $c > 0$, then $\Pr(X \geq c) \leq \frac{E(X)}{c}$.

Proof.
Let $c > 0$. Define

$$I = \begin{cases} 1 & \text{if } X \geq c \\ 0 & \text{otherwise.} \end{cases}$$

Since $X \geq 0$, $I \leq \frac{X}{c}$. Taking expectations of both side we find $E(I) \leq \frac{E(X)}{c}$. Now the result follows since $E(I) = \Pr(X \geq c)$ ■

Example 50.1
Suppose that a student's score on a test is a random variable with mean 75. Give an upper bound for the probability that a student's test score will exceed 85.

Solution.
Let X be the random variable denoting the student's score. Using Markov's inequality, we have

$$\Pr(X \geq 85) \leq \frac{E(X)}{85} = \frac{75}{85} \approx 0.882 \blacksquare$$

Example 50.2
If X is a non-negative random variable with $E(X) > 0$. Show that$\Pr(X \geq aE(X)) \leq \frac{1}{a}$ for all $a > 0$.

Solution.
The result follows by letting $c = aE(X)$ is Markov's inequality ■

Remark 50.1
Markov's inequality does not apply for negative random variable. To see this, let X be a random variable with range $\{-1000, 1000\}$. Suppose that $\Pr(X = -1000) = \Pr(X = 1000) = \frac{1}{2}$. Then $E(X) = 0$ and $\Pr(X \geq 1000) \neq 0$

Markov's bound gives us an upper bound on the probability that a random variable is large. It turns out, though, that there is a related result to get an upper bound on the probability that a random variable is small.

Proposition 50.2
Suppose that X is a random variable such that $X \leq M$ for some constant M. Then for all $x < M$ we have

$$\Pr(X \leq x) \leq \frac{M - E(X)}{M - x}.$$

Proof.
By applying Markov's inequality we find

$$\begin{aligned}\Pr(X \le x) =& \Pr(M - X \ge M - x)\\ \le& \frac{E(M-X)}{M-x} = \frac{M - E(X)}{M-x} \blacksquare\end{aligned}$$

Example 50.3
Let X denote the test score of a randomly chosen student, where the highest possible score is 100. Find an upper bound of $\Pr(X \le 50)$, given that $E(X) = 75$.

Solution.
By the previous proposition we find

$$\Pr(X \le 50) \le \frac{100-75}{100-50} = \frac{1}{2} \blacksquare$$

As a corollary of Proposition 50.1 we have

Proposition 50.3 *(Chebyshev's Inequality)*
If X is a random variable with finite mean μ and variance σ^2, then for any value $\epsilon > 0$,

$$\Pr(|X - \mu| \ge \epsilon) \le \frac{\sigma^2}{\epsilon^2}.$$

Proof.
Since $(X-\mu)^2 \ge 0$, by Markov's inequality we can write

$$\Pr((X-\mu)^2 \ge \epsilon^2) \le \frac{E[(X-\mu)^2]}{\epsilon^2}.$$

But $(X-\mu)^2 \ge \epsilon^2$ is equivalent to $|X - \mu| \ge \epsilon$ and this in turn is equivalent to

$$\Pr(|X-\mu| \ge \epsilon) \le \frac{E[(X-\mu)^2]}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2} \blacksquare$$

Example 50.4
Show that for any random variable the probability of a deviation from the mean of more than k standard deviations is less than or equal to $\frac{1}{k^2}$.

Solution.
This follows from Chebyshev's inequality by using $\epsilon = k\sigma$ ■

Example 50.5
Suppose X is the test score of a randomly selected student. We assume $0 \le X \le 100$, $E(X) = 70$, and $\sigma = 7$. Find an upper bound of $\Pr(X \ge 84)$ using first Markov's inequality and then Chebyshev's inequality.

Solution.
By using Markov's inequality we find

$$\Pr(X \ge 84) \le \frac{70}{84} = \frac{35}{42}$$

Now, using Chebyshev's inequality we find

$$\begin{aligned}\Pr(X \ge 84) =&\Pr(X - 70 \ge 14)\\ =&\Pr(X - E(X) \ge 2\sigma)\\ \le&\Pr(|X - E(X)| \ge 2\sigma) \le \frac{1}{4}\end{aligned}$$ ■

Example 50.6
The expected life of a certain battery is 240 hours.
(a) Let p be the probability that a battery will NOT last for 300 hours. What can you say about p?
(b) Assume now that the standard deviation of a battery's life is 30 hours. What can you say now about p?

Solution.
(a) Let X be the random variable representing the number of hours of the battery's life. Then by using Markov's inequality we find

$$p = \Pr(X < 300) = 1 - \Pr(X \ge 300) \ge 1 - \frac{240}{300} = 0.2$$

(b) By Chebyshev's inequality we find

$$p = \Pr(X < 300) = 1 - \Pr(X \ge 300) \ge 1 - \Pr(|X - 240| \ge 60) \ge 1 - \frac{900}{3600} = 0.75$$ ■

Example 50.7
You toss a fair coin n times. Assume that all tosses are independent. Let X denote the number of heads obtained in the n tosses.
(a) Compute (explicitly) the variance of X.
(b) Show that $\Pr(|X - E(X)| \geq \frac{n}{3}) \leq \frac{9}{4n}$.

Solution.
(a) For $1 \leq i \leq n$, let $X_i = 1$ if the i^{th} toss shows heads, and $X_i = 0$ otherwise. Thus, $X = X_1 + X_2 + \cdots + X_n$. Moreover, $E(X_i) = \frac{1}{2}$ and $E(X_i^2) = \frac{1}{2}$. Hence, $E(X) = \frac{n}{2}$ and

$$\begin{aligned} E(X^2) =& E\left[\left(\sum_{i=1}^n X_i\right)^2\right] = nE(X_1^2) + \sum\sum_{i \neq j} E(X_iX_j) \\ =& \frac{n}{2} + n(n-1)\frac{1}{4} = \frac{n(n+1)}{4} \end{aligned}$$

Hence, $\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{n(n+1)}{4} - \frac{n^2}{4} = \frac{n}{4}$.
(b) We apply Chebychev's inequality:

$$\Pr(|X - E(X)| \geq \frac{n}{3}) \leq \frac{\text{Var}(X)}{(n/3)^2} = \frac{9}{4n} \blacksquare$$

When does a random variable, X, have zero variance? It turns out that this happens when the random variable never deviates from the mean. The following theorem characterizes the structure of a random variable whose variance is zero.

Proposition 50.4
If X is a random variable with zero variance, then X must be constant with probability equals to 1.

Proof.
First we show that if $X \geq 0$ and $E(X) = 0$ then $X = 0$ and $\Pr(X = 0) = 1$. Since $E(X) = 0$, by Markov's inequality $\Pr(X \geq c) = 0$ for all $c > 0$. But

$$\Pr(X > 0) = P\left(\bigcup_{n=1}^{\infty}(X > \frac{1}{n})\right) \leq \sum_{n=1}^{\infty} \Pr(X > \frac{1}{n}) = 0.$$

Hence, $\Pr(X > 0) = 0$. Since $X \geq 0$, $1 = \Pr(X \geq 0) = \Pr(X = 0) + \Pr(X > 0) = \Pr(X = 0)$.
Now, suppose that $Var(X) = 0$. Since $(X - E(X))^2 \geq 0$ and $Var(X) = E((X - E(X))^2)$, by the above result we have $\Pr(X - E(X) = 0) = 1$. That is, $\Pr(X = E(X)) = 1$ ∎

One of the most well known and useful results of probability theory is the following theorem, known as the **weak law of large numbers**.

Theorem 50.1
Let $X_1, X_2, \cdots, X_n$ be a sequence of independent random variables with common mean μ and finite common variance σ^2. Then for any $\epsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\tfrac{X_1+X_2+\cdots+X_n}{n} - \mu\right| \geq \epsilon\right\} = 0$$

or equivalently

$$\lim_{n\to\infty} P\left(\left|\frac{X_1+X_2+\cdots+X_n}{n} - \mu\right| < \epsilon\right) = 1$$

Proof.
Since

$$E\left[\tfrac{X_1+X_2+\cdots+X_n}{n}\right] = \mu \text{ and } \mathrm{Var}\left(\tfrac{X_1+X_2+\cdots+X_n}{n}\right) = \tfrac{\sigma^2}{n}$$

by Chebyshev's inequality we find

$$0 \leq P\left\{\left|\frac{X_1+X_2+\cdots+X_n}{n} - \mu\right| \geq \epsilon\right\} \leq \frac{\sigma^2}{n\epsilon^2}$$

and the result follows by letting $n \to \infty$ ∎

The above theorem says that for large n, $\frac{X_1+X_2+\cdots+X_n}{n} - \mu$ is small with high probability. Also, it says that the distribution of the sample average becomes concentrated near μ as $n \to \infty$.
Let A be an event with probability p. Repeat the experiment n times. Let X_i be 1 if the event occurs and 0 otherwise. Then $S_n = \frac{X_1+X_2+\cdots+X_n}{n}$ is the number of occurrence of A in n trials and $\mu = E(X_i) = p$. By the weak law of large numbers we have

$$\lim_{n\to\infty} P\left(|S_n - \mu| < \epsilon\right) = 1$$

The above statement says that, in a large number of repetitions of a Bernoulli experiment, we can expect the proportion of times the event will occur to be near $p = \Pr(A)$. This agrees with the definition of probability that we introduced in Section 6.
The Weak Law of Large Numbers was given for a sequence of pairwise independent random variables with the same mean and variance. We can generalize the Law to sequences of pairwise independent random variables, possibly with different means and variances, as long as their variances are bounded by some constant.

Example 50.8
Let $X_1, X_2, \cdots$ be pairwise independent random variables such that $\text{Var}(X_i) \le b$ for some constant $b > 0$ and for all $1 \le i \le n$. Let

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

and

$$\mu_n = E(S_n).$$

Show that, for every $\epsilon > 0$ we have

$$\Pr(|S_n - \mu_n| > \epsilon) \le \frac{b}{\epsilon^2} \cdot \frac{1}{n}$$

and consequently

$$\lim_{n\to\infty} \Pr(|S_n - \mu_n| \le \epsilon) = 1$$

Solution.
Since $E(S_n) = \mu_n$ and $\text{Var}(S_n) = \frac{\text{Var}(X_1)+\text{Var}(X_2)+\cdots+\text{Var}(X_n)}{n^2} \le \frac{bn}{n^2} = \frac{b}{n}$, by Chebyshev's inequality we find

$$0 \le P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu_n\right| \ge \epsilon\right\} \le \frac{1}{\epsilon^2} \cdot \frac{b}{n}.$$

Now,

$$1 \ge \Pr(|S_n - \mu_n| \le \epsilon) = 1 - \Pr(|S_n - \mu_n| > \epsilon) \ge 1 - \frac{b}{\epsilon^2} \cdot \frac{1}{n}$$

By letting $n \to \infty$ we conclude that

$$\lim_{n\to\infty} \Pr(|S_n - \mu_n| \le \epsilon) = 1 \blacksquare$$

50.2 The Strong Law of Large Numbers

Recall the weak law of large numbers:

$$\lim_{n\to\infty} \Pr(|S_n - \mu| < \epsilon) = 1$$

where the X_i's are independent identically distributed random variables and $S_n = \frac{X_1+X_2+\cdots+X_n}{n}$. This type of convergence is referred to as **convergence in probability.** Unfortunately, this form of convergence does not assure convergence of individual realizations. In other words, for any given elementary event $x \in S$, we have no assurance that $\lim_{n\to\infty} S_n(x) = \mu$. Fortunately, however, there is a stronger version of the law of large numbers that does assure convergence for individual realizations.

Theorem 50.2 *(Strong law of large numbers)*
Let $\{X_n\}_{n\geq 1}$ be a sequence of independent random variables with finite mean $\mu = E(X_i)$ and $K = E(X_i^4) < \infty$. Then

$$P\left(\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right) = 1.$$

Proof.
We first consider the case $\mu = 0$ and let $T_n = X_1 + X_2 + \cdots + X_n$. Then $E(T_n^4) = E[(X_1 + X_2 + \cdots + X_n)(X_1 + X_2 + \cdots + X_n)(X_1 + X_2 + \cdots + X_n)(X_1 + X_2 + \cdots + X_n)]$
When expanding the product on the right side using the multinomial theorem the resulting expression contains terms of the form

$$X_i^4, \quad X_i^3X_j, \quad X_i^2X_j^2, \quad X_i^2X_jX_k \quad \text{and} \quad X_iX_jX_kX_l$$

with $i \neq j \neq k \neq l$. Now recalling that $\mu = 0$ and using the fact that the random variables are independent we find

$$\begin{aligned}
E(X_i^3X_j) =& E(X_i^3)E(X_j) = 0\\
E(X_i^2X_jX_k) =& E(X_i^2)E(X_j)E(X_k) = 0\\
E(X_iX_jX_kX_l) =& E(X_i)E(X_j)E(X_k)E(X_l) = 0
\end{aligned}$$

Next, there are n terms of the form X_i^4 and for each $i \neq j$ the coefficient of $X_i^2X_j^2$ according to the multinomial theorem is

$$\frac{4!}{2!2!} = 6.$$

But there are ${}_nC_2 = \frac{n(n-1)}{2}$ different pairs of indices $i \neq j$. Thus, by taking the expectation term by term of the expansion we obtain

$$E(T_n^4) = nE(X_i^4) + 3n(n-1)E(X_i^2)E(X_j^2)$$

where in the last equality we made use of the independence assumption. Now, from the definition of the variance we find

$$0 \leq Var(X_i^2) = E(X_i^4) - (E(X_i^2))^2$$

and this implies that

$$(E(X_i^2))^2 \leq E(X_i^4) = K.$$

It follows that

$$E(T_n^4) \leq nK + 3n(n-1)K$$

which implies that

$$E\left[\frac{T_n^4}{n^4}\right] \leq \frac{K}{n^3} + \frac{3K}{n^2} \leq \frac{4K}{n^2}$$

Therefore.

$$E\left[\sum_{n=1}^{\infty} \frac{T_n^4}{n^4}\right] \leq 4K \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty \tag{50.1}$$

Now,

$$P\left[\sum_{n=1}^{\infty} \frac{T_n^4}{n^4} < \infty\right] + P\left[\sum_{n=1}^{\infty} \frac{T_n^4}{n^4} = \infty\right] = 1.$$

If $P\left[\sum_{n=1}^{\infty} \frac{T_n^4}{n^4} = \infty\right] > 0$ then at some value in the range of the random variable $\sum_{n=1}^{\infty} \frac{T_n^4}{n^4}$ the sum $\sum_{n=1}^{\infty} \frac{T_n^4}{n^4}$ is infinite and so its expected value is infinite which contradicts (50.1). Hence, $P\left[\sum_{n=1}^{\infty} \frac{T_n^4}{n^4} = \infty\right] = 0$ and therefore

$$P\left[\sum_{n=1}^{\infty} \frac{T_n^4}{n^4} < \infty\right] = 1.$$

But the convergence of a series implies that its nth term goes to 0. Hence,

$$P\left[\lim_{n \to \infty} \frac{T_n^4}{n^4} = 0\right] = 1$$

Since $\frac{T_n^4}{n^4} = \left(\frac{T_n}{n}\right)^4 = S_n^4$ the last result implies that

$$P\left[\lim_{n\to\infty} S_n = 0\right] = 1$$

which proves the result for $\mu = 0$.
Now, if $\mu \neq 0$, we can apply the preceding argument to the random variables $X_i - \mu$ to obtain

$$P\left[\lim_{n\to\infty} \sum_{i=1}^{n} \frac{(X_i - \mu)}{n} = 0\right] = 1$$

or equivalently

$$P\left[\lim_{n\to\infty} \sum_{i=1}^{n} \frac{X_i}{n} = \mu\right] = 1$$

which proves the theorem ■

As an application of this theorem, suppose that a sequence of independent trials of some experiment is performed. Let E be a fixed event of the experiment and let $\Pr(E)$ denote the probability that E occurs on any particular trial. Define

$$X_i = \begin{cases} 1 & \text{if } E \text{ occurs on the ith trial} \\ 0 & \text{otherwise} \end{cases}$$

By the Strong Law of Large Numbers we have

$$P\left[\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = E(X) = \Pr(E)\right] = 1$$

Since $X_1 + X_2 + \cdots + X_n$ represents the number of times the event E occurs in the first n trials, the above result says that the limiting proportion of times that the event occurs is just $\Pr(E)$. This justifies our definition of $\Pr(E)$ that we introduced in Section 6,i.e.,

$$\Pr(E) = \lim_{n\to\infty} \frac{n(E)}{n}$$

where $n(E)$ denotes the number of times in the first n repetitions of the experiment that the event E occurs.

To clarify the somewhat subtle difference between the Weak and Strong Laws of Large Numbers, we will construct an example of a sequence $X_1, X_2, \cdots$ of mutually independent random variables that satisfies the Weak Law of Large but not the Strong Law.

Example 50.9
Let $X_1, X_2, \cdots$ be the sequence of mutually independent random variables such that $X_1 = X_2 = 0$, and for each positive integer $i > 2$

$$\Pr(X_i = i) = \tfrac{1}{2i\ln i}, \quad \Pr(X_i = -i) = \tfrac{1}{2i\ln i}, \quad \Pr(X_i = 0) = 1 - \tfrac{1}{i\ln i}$$

Note that $E(X_i) = 0$ for all i. Let $T_n = X_1 + X_2 + \cdots + X_n$ and $S_n = \frac{T_n}{n}$.

(a) Show that $\text{Var}(T_n) \le \frac{n^2}{\ln n}$.
(b) Show that the sequence $X_1, X_2, \cdots$ satisfies the Weak Law of Large Numbers, i.e., prove that for any $\epsilon > 0$

$$\lim_{n\to\infty} \Pr(|S_n| \ge \epsilon) = 0.$$

(c) Let $A_1, A_2, \cdots$ be any infinite sequence of mutually independent events such that

$$\sum_{i=1}^{\infty} \Pr(A_i) = \infty.$$

Prove that

$$\Pr(\text{infinitely many } A_i \text{ occur}) = 1$$

(d) Show that $\sum_{i=1}^{\infty} \Pr(|X_i| \ge i) = \infty$.
(e) Conclude that

$$\lim_{n\to\infty} \Pr(S_n = \mu) = 0$$

and hence that the Strong Law of Large Numbers completely fails for the sequence $X_1, X_2, \cdots$

Solution.
(a) We have

$$\begin{aligned}\mathrm{Var}(T_n) =&\mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n)\\ =&0 + 0 + \sum_{i=3}^{n}(E(X_i^2) - (E(X_i))^2)\\ =&\sum_{i=3}^{n} i^2\frac{1}{i\ln i}\\ =&\sum_{i=3}^{n}\frac{i}{\ln i}\end{aligned}$$

Now, if we let $f(x) = \frac{x}{\ln x}$ then $f'(x) = \frac{1}{\ln x}\left(1 - \frac{1}{\ln x}\right) > 0$ for $x > e$ so that $f(x)$ is increasing for $x > e$. It follows that $\frac{n}{\ln n} \geq \frac{i}{\ln i}$ for $3 \leq i \leq n$. Furthermore,

$$\frac{n^2}{\ln n} = \sum_{i=1}^{n}\frac{n}{\ln n} \geq \sum_{i=3}^{n}\frac{i}{\ln i}$$

Hence,

$$\mathrm{Var}(T_n) \leq \frac{n^2}{\ln n}.$$

(b) We have

$$\begin{aligned}\Pr(|S_n| \geq \epsilon) =&\Pr(|S_n - 0| \geq \epsilon)\\ \leq&\mathrm{Var}(S_n)\cdot\frac{1}{\epsilon^2} \quad \text{Chebyshev's inequality}\\ =&\frac{\mathrm{Var}(T_n)}{n^2}\cdot\frac{1}{\epsilon^2}\\ \leq&\frac{1}{\epsilon^2\ln n}\end{aligned}$$

which goes to zero as n goes to infinity.
(c) Let $T_{r,n} = \sum_{i=r}^{n} I_{A_i}$ the number of events A_i with $r \leq i \leq n$ that occur. Then

$$\lim_{n\to\infty} E(T_{r,n}) = \lim_{n\to\infty}\sum_{i=r}^{n} E(I_{A_i}) = \lim_{n\to\infty}\sum_{i=r}^{n}\Pr(A_i) = \infty$$

Since $e^x \to 0$ as $x \to -\infty$ we conclude that $e^{-E(T_{r,n})} \to 0$ as $n \to \infty$.
Now, let K_r be the event that no A_i with $i \geq r$ occurs. Also, let $K_{r,n}$ be the

event that no A_i with $r \leq i \leq n$ occurs. Finally, let K be the event that only finitely many $A_i's$. occurs. We must prove that $\Pr(K) = 0$. We first show that

$$\Pr(K_{r,n}) \leq e^{-E(T_{r,n})}.$$

We remind the reader of the inequality $1 + x \leq e^x$. We have

$$\begin{aligned}
\Pr(K_{r,n}) =&\Pr(T_{r,n} = 0) = P[(A_r \cup A_{r+1} \cup \cdots \cup A_n)^c] \\
=&P[A_r^c \cap A_{r+1}^c \cap \cdots \cap A_n^c] \\
=&\prod_{i=r}^{n} \Pr(A_i^c) \\
=&\prod_{i=r}^{n} [1 - \Pr(A_i)] \\
\leq&\prod_{i=r}^{n} e^{-\Pr(A_i)} \\
=&e^{-\sum_{i=r}^{n} \Pr(A_i)} \\
=&e^{-\sum_{i=r}^{n} E(I_{A_i})} \\
=&e^{-E(T_{r,n})}
\end{aligned}$$

Now, since $K_r \subset K_{r,n}$ we conclude that $0 \leq \Pr(K_r) \leq \Pr(K_{r,n}) \leq e^{-E(T_{r,n})} \to 0$ as $n \to \infty$. Hence, $\Pr(K_r) = 0$ for all $r \leq n$.
Now note that $K = \cup_r K_r$ so by Boole's inequality (See Proposition 46.3), $0 \leq \Pr(K) \leq \sum_r K_r = 0$. That is, $\Pr(K) = 0$. Hence the probability that infinitely many A_i's occurs is 1.
(d) We have that $\Pr(|X_i| \geq i) = \frac{1}{i \ln i}$. Thus,

$$\begin{aligned}
\sum_{i=1}^{n} \Pr(|X_i| \geq i) =&0 + 0 + \sum_{i=3}^{n} \frac{1}{i \ln i} \\
\geq& \int_3^n \frac{dx}{x \ln x} \\
=& \ln \ln n - \ln \ln 3
\end{aligned}$$

and this last term approaches infinity as n approaches infinity.
(e) By parts (c) and (d), the probability that $|X_i| \geq i$ for infinitely many i is 1. But if $|X_i| \geq i$ for infinitely many i then from the definition of limit

$\lim_{n\to\infty} \frac{X_n}{n} \neq 0$. Hence,

$$\Pr(\lim_{n\to\infty} \frac{X_n}{n} \neq 0) = 1$$

which means

$$\Pr(\lim_{n\to\infty} \frac{X_n}{n} = 0) = 0$$

Now note that

$$\frac{X_n}{n} = S_n - \frac{n-1}{n} S_{n-1}$$

so that if $\lim_{n\to\infty} S_n = 0$ then $\lim_{n\to\infty} \frac{X_n}{n} = 0$. This implies that

$$\Pr(\lim_{n\to\infty} S_n = 0) \leq \Pr(\lim_{n\to\infty} \frac{X_n}{n} = 0)$$

That is,

$$\Pr(\lim_{n\to\infty} S_n = 0) = 0$$

and this violates the Stong Law of Large numbers ■

Practice Problems

Problem 50.1
Let $\epsilon > 0$. Let X be a discrete random variable with range $\{-\epsilon, \epsilon\}$ and pmf given by $\Pr(-\epsilon) = \frac{1}{2}$ and $\Pr(\epsilon) = \frac{1}{2}$ and 0 otherwise. Show that the inequality in Chebyshev's inequality is in fact an equality.

Problem 50.2
Let $X_1, X_2, \cdots, X_n$ be a Bernoulli trials process with probability 0.3 for success and 0.7 for failure. Let $X_i = 1$ if the ith outcome is a success and 0 otherwise. Find n so that $P\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \geq 0.1\right) \leq 0.21$, where $S_n = X_1 + X_2 + \cdots + X_n$.

Problem 50.3
Suppose that X is uniformly distributed on $[0, 12]$. Find an upper bound for the probability that a sample from X lands more than 1.5 standard deviation from the mean.

Problem 50.4
Let X be a random variable with $E(X) = 10^3$. Find an upper bound for the probability that X is at least 10^4.

Problem 50.5
Suppose that X is a random variable with mean and variance both equal to 20. What can be said about $\Pr(0 < X < 40)$?

Problem 50.6
Let $X_1, X_2, \cdots, X_{20}$ be independent Poisson random variables with mean 1. Use Markov's inequality to obtain a bound on

$$\Pr(X_1 + X_2 + \cdots + X_{20} > 15).$$

Problem 50.7
Suppose that the test score of a randomly chosen student has mean 75 and variance 25. What can be said about the probability that a student will score between 65 and 85?

Problem 50.8
Let $M_X(t) = E(e^{tX})$ be the moment generating function of a random variable X. Show that

$$\Pr(X \geq \epsilon) \leq e^{-\epsilon t} M_X(t),\ t > 0.$$

Problem 50.9
The expected weekly sale of computers by a store is 50. Suppose that on Monday the store has 75 computers. What can be said about the probability of not having enough computers left by the end of the week?

Problem 50.10
Let X be a random variable with mean $\frac{1}{2}$ and variance 25×10^{-7}. What can you say about $\Pr(0.475 \leq X \leq 0.525)$?

Problem 50.11
Let $X \geq 0$ be a random variable with mean μ. Show that $\Pr(X \geq 2\mu) \leq \frac{1}{2}$.

Problem 50.12
The expected daily production of a DVD factory is 100 DVDs.
(a) Find an upper bound to the probability that the factory's production will be more than 120 in a day.
(b) Suppose that the variance of the daily production is known to be 5. Find a lower bound to the probability that the factory's production will be between 70 and 130 in a day.

Problem 50.13
A biased coin comes up heads 30% of the time. The coin is tossed 400 times. Let X be the number of heads in the 400 tossings. Use Chebyshev's inequality to bound the probability that X is between 100 and 140.

Problem 50.14
Let $X_1, \cdots, X_n$ be independent random variables, each with probability density function:

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Show that $\overline{X} = \frac{\sum_{i=1}^n X_i}{n}$ converges in probability to a constant as $n \to \infty$ and find that constant.

Problem 50.15
Let $X_1, \cdots, X_n$ be independent and identically distributed Uniform(0,1). Let Y_n be the minimum of $X_1, \cdots, X_n$.
(a) Find the cumulative distribution of Y_n
(b) Show that Y_n converges in probability to 0 by showing that for arbitrary $\epsilon > 0$

$$\lim_{n \to \infty} \Pr(|Y_n - 0| \leq \epsilon) = 1.$$

51 The Central Limit Theorem

The central limit theorem is one of the most remarkable theorems among the limit theorems. This theorem says that the sum of a large number of independent identically distributed random variables is well-approximated by a normal random variable.
We first need a technical result.

Theorem 51.1
Let $Z_1, Z_2, \cdots$ be a sequence of random variables having distribution functions F_{Z_n} and moment generating functions $M_{Z_n}, n \geq 1$. Let Z be a random variable having distribution F_Z and moment generating function M_Z. If $M_{Z_n}(t) \to M_Z(t)$ as $n \to \infty$ and for all t, then $F_{Z_n} \to F_Z$ for all t at which $F_Z(t)$ is continuous.

With the above theorem, we can prove the **central limit theorem**.

Theorem 51.2
Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables, each with mean μ and variance σ^2. Then,

$$P\left(\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+X_2+\cdots+X_n}{n}-\mu\right) \leq a\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-\frac{x^2}{2}}dx$$

as $n \to \infty$.

The Central Limit Theorem says that regardless of the underlying distribution of the variables X_i, so long as they are independent, the distribution of $\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+X_2+\cdots+X_n}{n}-\mu\right)$ converges to the same, normal, distribution.

Proof.
We prove the theorem under the assumption that $E(e^{tX_i})$ is finite in a neighborhood of 0. In particular, we show that

$$M_{\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+X_2+\cdots+X_n}{n}-\mu\right)}(x) \to e^{\frac{x^2}{2}}$$

where $e^{\frac{x^2}{2}}$ is the moment generating function of the standard normal distribution.

Now, using the properties of moment generating functions we can write

$$\begin{aligned}
M_{\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+X_2+\cdots+X_n}{n}-\mu\right)}(x) =&M_{\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{X_i-\mu}{\sigma}}(x)\\
=&M_{\sum_{i=1}^n \frac{X_i-\mu}{\sigma}}\left(\frac{x}{\sqrt{n}}\right)\\
=&\prod_{i=1}^n M_{\frac{X_i-\mu}{\sigma}}\left(\frac{x}{\sqrt{n}}\right)\\
=&\left(M_{\frac{X_1-\mu}{\sigma}}\left(\frac{x}{\sqrt{n}}\right)\right)^n\\
=&\left(M_Y\left(\frac{x}{\sqrt{n}}\right)\right)^n
\end{aligned}$$

where

$$Y=\frac{X_1-\mu}{\sigma}.$$

Now expand $M_Y(x/\sqrt{n})$ in a Taylor series around 0 as follows

$$M_Y\left(\frac{x}{\sqrt{n}}\right)=M_Y(0)+M_Y'(0)\left(\frac{x}{\sqrt{n}}\right)+\frac{1}{2}\frac{x^2}{n}M_Y''(0)+R\left(\frac{x}{\sqrt{n}}\right),$$

where

$$\tfrac{n}{x^2}R\left(\tfrac{x}{\sqrt{n}}\right)\to 0 \text{ as } x\to 0$$

But

$$E(Y)=E\left[\frac{X_1-\mu}{\sigma}\right]=0$$

and

$$E(Y^2)=E\left[\left(\frac{X_1-\mu}{\sigma}\right)^2\right]=\frac{Var(X_1)}{\sigma^2}=1.$$

By Proposition 49.1 we obtain $M_Y(0)=1$, $M_Y'(0)=E(Y)=0$, and $M_Y''(0)=E(Y^2)=1$. Hence,

$$M_Y\left(\frac{x}{\sqrt{n}}\right)=1+\frac{1}{2}\frac{x^2}{n}+R\left(\frac{x}{\sqrt{n}}\right)$$

and so

$$\begin{aligned}\lim_{n\to\infty}\left(M_Y\left(\frac{x}{\sqrt{n}}\right)\right)^n =& \lim_{n\to\infty}\left(1+\frac{1}{2}\frac{x^2}{n}+R\left(\frac{x}{\sqrt{n}}\right)\right)^n\\ =& \lim_{n\to\infty}\left(1+\frac{1}{n}\left(\frac{x^2}{2}+nR\left(\frac{x}{\sqrt{n}}\right)\right)\right)^n\end{aligned}$$

But

$$nR\left(\frac{x}{\sqrt{n}}\right)=x^2\frac{n}{x^2}R\left(\frac{x}{\sqrt{n}}\right)\to 0$$

as $n\to\infty$. Hence,

$$\lim_{n\to\infty}\left(M_Y\left(\frac{x}{\sqrt{n}}\right)\right)^n=e^{\frac{x^2}{2}}.$$

Now the result follows from Theorem 51.1 with $Z_n=\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+X_2+\cdots+X_n}{n}-\mu\right)$, Z standard normal distribution,

$$F_{Z_n}(a)=P\left(\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+X_2+\cdots+X_n}{n}-\mu\right)\le a\right)$$

and

$$F_Z(a)=\Phi(a)=\int_{-\infty}^{a}e^{-\frac{x^2}{2}}dx \blacksquare$$

The central limit theorem suggests approximating the random variable

$$\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+X_2+\cdots+X_n}{n}-\mu\right)$$

with a standard normal random variable. This implies that the sample mean has approximately a normal distribution with mean μ and variance $\frac{\sigma^2}{n}$.
Also, a sum of n independent and identically distributed random variables with common mean μ and variance σ^2 can be approximated by a normal distribution with mean $n\mu$ and variance $n\sigma^2$.

Example 51.1 ‡
In an analysis of healthcare data, ages have been rounded to the nearest multiple of 5 years. The difference between the true age and the rounded age is assumed to be uniformly distributed on the interval from −2.5 years to 2.5 years. The healthcare data are based on a random sample of 48 people.
What is the approximate probability that the mean of the rounded ages is within 0.25 years of the mean of the true ages?

Solution.
Let X denote the difference between true and reported age. We are given X is uniformly distributed on $(-2.5, 2.5)$. That is, X has pdf $f(x) = 1/5, -2.5 < x < 2.5$. It follows that $E(X) = 0$ and

$$\sigma_X^2 = E(X^2) = \int_{-2.5}^{2.5} \frac{x^2}{5} dx \approx 2.083$$

so that $SD(X) = \sqrt{2.083} \approx 1.443$.
Now $\overline{X}_{48}$ the difference between the means of the true and rounded ages, has a distribution that is approximately normal with mean 0 and standard deviation $\frac{1.443}{\sqrt{48}} \approx 0.2083$. Therefore,

$$\begin{aligned} P\left(-\frac{1}{4} \le \overline{X}_{48} \le \frac{1}{4}\right) =& P\left(\frac{-0.25}{0.2083} \le \frac{X_{48}}{0.2083} \le \frac{0.25}{0.2083}\right) \\ =& \Pr(-1.2 \le Z \le 1.2) = 2\Phi(1.2) - 1 \approx 2(0.8849) - 1 = 0.77 \blacksquare \end{aligned}$$

Example 51.2
Let $X_i,\ i = 1, 2, \cdots, 48$ be independent random variables that are uniformly distributed on the interval $[-0.5, 0.5]$. Find the approximate probability $\Pr(|\overline{X}| \le 0.05)$, where $\overline{X}$ is the arithmetic average of the $X_i's$.

Solution.
Since each X_i is uniformly distributed on $[-0.5, 0.5]$, its mean is $\mu = 0$ and its variance is $\sigma^2 = \int_{-0.5}^{0.5} x^2 dx = \left.\frac{x^3}{3}\right]_{-0.5}^{0.5} = \frac{1}{12}$. By the Central Limit Theorem, $\overline{X}$ has approximate distribution $N(\mu, \frac{\sigma^2}{n}) = N(0, \frac{1}{24^2})$. Thus $24\overline{X}$ is approximately standard normal, so

$$\begin{aligned} \Pr(|\overline{X}| \le 0.05) \approx& \Pr(24 \cdot (-0.05) \le 24\overline{X} \le 24 \cdot (0.05)) \\ =& \Phi(1.2) - \Phi(-1.2) = 2\Phi(1.2) - 1 = 0.7698 \blacksquare \end{aligned}$$

Example 51.3
Let X_1, X_2, X_3, X_4 be a random sample of size 4 from a normal distribution with mean 2 and variance 10, and let $\overline{X}$ be the sample mean. Determine a such that $\Pr(\overline{X} \le a) = 0.90$.

Solution.
The sample mean $\overline{X}$ is normal with mean $\mu = 2$ and variance $\frac{\sigma^2}{n} = \frac{10}{4} = 2.5$, and standard deviation $\sqrt{2.5} \approx 1.58$, so

$$0.90 = \Pr(\overline{X} \leq a) = P\left(\frac{\overline{X} - 2}{1.58} < \frac{a-2}{1.58}\right) = \Phi\left(\frac{a-2}{1.58}\right).$$

From the normal table, we get $\frac{a-2}{1.58} = 1.28$, so $a = 4.02$ ■

Example 51.4 [6]
Assume that the weights of individuals are independent and normally distributed with a mean of 160 pounds and a standard deviation of 30 pounds. Suppose that 25 people squeeze into an elevator that is designed to hold 4300 pounds.
(a) What is the probability that the load (total weight) exceeds the design limit?
(b) What design limit is exceeded by 25 occupants with probability 0.001?

Solution.
(a) Let X be an individual's weight. Then, X has a normal distribution with $\mu = 160$ pounds and $\sigma = 30$ pounds. Let $Y = X_1 + X_2 + \cdots + X_{25}$, where X_i denotes the i^{th} person's weight. Then, Y has a normal distribution with $E(Y) = 25E(X) = 25 \cdot (160) = 4000$ pounds and $\text{Var}(Y) = 25\text{Var}(X) = 25 \cdot (900) = 22500$. Now, the desired probability is

$$\begin{aligned}\Pr(Y > 4300) =& P\left(\frac{Y - 4000}{\sqrt{22500}} > \frac{4300 - 4000}{\sqrt{22500}}\right)\\ =& \Pr(Z > 2) = 1 - \Pr(Z \leq 2)\\ =& 1 - 0.9772 = 0.0228\end{aligned}$$

(b) We want to find x such that $\Pr(Y > x) = 0.001$. Note that

$$\begin{aligned}\Pr(Y > x) =& P\left(\frac{Y - 4000}{\sqrt{22500}} > \frac{x - 4000}{\sqrt{22500}}\right)\\ =& P\left(Z > \frac{x - 4000}{\sqrt{22500}}\right) = 0.01\end{aligned}$$

[6] *Applied Statisitcs and Probability for Engineers* by Montgomery and Tunger

It is equivalent to $P\left(Z \leq \frac{x-4000}{\sqrt{22500}}\right) = 0.999$. From the normal Table we find $\Pr(Z \leq 3.09) = 0.999$. So $(x - 4000)/150 = 3.09$. Solving for x we find $x \approx 4463.5$ pounds ■

Practice Problems

Problem 51.1
Letter envelopes are packaged in boxes of 100. It is known that, on average, the envelopes weigh 1 ounce, with a standard deviation of 0.05 ounces. What is the probability that 1 box of envelopes weighs more than 100.4 ounces?

Problem 51.2
In the SunBelt Conference men basketball league, the standard deviation in the distribution of players' height is 2 inches. A random group of 25 players are selected and their heights are measured. Estimate the probability that the average height of the players in this sample is within 1 inch of the conference average height.

Problem 51.3
A radio battery manufacturer claims that the lifespan of its batteries has a mean of 54 days and a standard deviation of 6 days. A random sample of 50 batteries were picked for testing. Assuming the manufacturer's claims are true, what is the probability that the sample has a mean lifetime of less than 52 days?

Problem 51.4
If 10 fair dice are rolled, find the approximate probability that the sum obtained is between 30 and 40, inclusive.

Problem 51.5
Let $X_i, i = 1, 2, \cdots, 10$ be independent random variables each uniformly distributed over (0,1). Calculate an approximation to $\Pr(\sum_{i=1}^{10} X_i > 6)$.

Problem 51.6
Suppose that $X_i, i = 1, \cdots, 100$ are exponentially distributed random variables with parameter $\lambda = \frac{1}{1000}$. Let $\overline{X} = \frac{\sum_{i=1}^{100} X_i}{100}$. Approximate $\Pr(950 \leq \overline{X} \leq 1050)$.

Problem 51.7
A baseball team plays 100 independent games. It is found that the probability of winning a game is 0.8. Estimate the probability that team wins at least 90 games.

Problem 51.8
A small auto insurance company has 10,000 automobile policyholders. It has found that the expected yearly claim per policyholder is \$240 with a standard deviation of \$800. Estimate the probability that the total yearly claim exceeds \$2.7 million.

Problem 51.9
Let $X_1, X_2, \cdots, X_n$ be n independent random variables each with mean 100 and standard deviation 30. Let X be the sum of these random variables. Find n such that $\Pr(X > 2000) \geq 0.95$.

Problem 51.10 ‡
A charity receives 2025 contributions. Contributions are assumed to be independent and identically distributed with mean 3125 and standard deviation 250.
Calculate the approximate 90th percentile for the distribution of the total contributions received.

Problem 51.11 ‡
An insurance company issues 1250 vision care insurance policies. The number of claims filed by a policyholder under a vision care insurance policy during one year is a Poisson random variable with mean 2. Assume the numbers of claims filed by distinct policyholders are independent of one another.
What is the approximate probability that there is a total of between 2450 and 2600 claims during a one-year period?

Problem 51.12 ‡
A company manufactures a brand of light bulb with a lifetime in months that is normally distributed with mean 3 and variance 1 . A consumer buys a number of these bulbs with the intention of replacing them successively as they burn out. The light bulbs have independent lifetimes.
What is the smallest number of bulbs to be purchased so that the succession of light bulbs produces light for at least 40 months with probability at least 0.9772?

Problem 51.13 ‡
Let X and Y be the number of hours that a randomly selected person watches movies and sporting events, respectively, during a three-month period. The following information is known about X and Y :

E(X) =	50
E(Y) =	20
Var(X) =	50
Var(Y) =	30
Cov (X,Y) =	10

One hundred people are randomly selected and observed for these three months. Let T be the total number of hours that these one hundred people watch movies or sporting events during this three-month period.
Approximate the value of $\Pr(T < 7100)$.

Problem 51.14 ‡
The total claim amount for a health insurance policy follows a distribution with density function

$$f(x) = \begin{cases} \frac{1}{1000}e^{-\frac{x}{1000}} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

The premium for the policy is set at 100 over the expected total claim amount. If 100 policies are sold, what is the approximate probability that the insurance company will have claims exceeding the premiums collected?

Problem 51.15 ‡
A city has just added 100 new female recruits to its police force. The city will provide a pension to each new hire who remains with the force until retirement. In addition, if the new hire is married at the time of her retirement, a second pension will be provided for her husband. A consulting actuary makes the following assumptions:

(i) Each new recruit has a 0.4 probability of remaining with the police force until retirement.
(ii) Given that a new recruit reaches retirement with the police force, the probability that she is not married at the time of retirement is 0.25.
(iii) The number of pensions that the city will provide on behalf of each new hire is independent of the number of pensions it will provide on behalf of any other new hire.
Determine the probability that the city will provide at most 90 pensions to the 100 new hires and their husbands.

Problem 51.16
(a) Give the approximate sampling distribution for the following quantity based on random samples of independent observations:

$$\overline{X} = \frac{\sum_{i=1}^{100} X_i}{100}, \quad E(X_i) = 100, \quad \mathrm{Var}(X_i) = 400.$$

(b) What is the approximate probability the sample mean will be between 96 and 104?

Problem 51.17
A biased coin comes up heads 30% of the time. The coin is tossed 400 times. Let X be the number of heads in the 400 tossings.
(a) Use Chebyshev's inequality to bound the probability that X is between 100 and 140.
(b) Use normal approximation to compute the probability that X is between 100 and 140.

52 More Useful Probabilistic Inequalities

The importance of the Markov's and Chebyshev's inequalities is that they enable us to derive bounds on probabilities when only the mean, or both the mean and the variance, of the probability distribution are known. In this section, we establish more probability bounds.
The following result gives a tighter bound in Chebyshev's inequality.

Proposition 52.1
Let X be a random variable with mean μ and finite variance σ^2. Then for any $a > 0$

$$\Pr(X \geq \mu + a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

and

$$\Pr(X \leq \mu - a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

Proof.
Without loss of generality we assume that $\mu = 0$. Then for any $b > 0$ we have

$$\begin{aligned}\Pr(X \geq a) =& \Pr(X + b \geq a + b)\\ \leq& \Pr((X+b)^2 \geq (a+b)^2)\\ \leq& \frac{E[(X+b)^2]}{(a+b)^2}\\ =& \frac{\sigma^2 + b^2}{(a+b)^2}\\ =& \frac{\alpha + t^2}{(1+t)^2} = g(t)\end{aligned}$$

where

$$\alpha = \tfrac{\sigma^2}{a^2} \text{ and } t = \tfrac{b}{a}.$$

Since

$$g'(t) = 2\frac{t^2 + (1-\alpha)t - \alpha}{(1+t)^4}$$

we find $g'(t) = 0$ when $t = \alpha$. Since $g''(t) = 2(2t + 1 - \alpha)(1+t)^{-4} - 8(t^2 + (1-\alpha)t - \alpha)(1+t)^{-5}$ we find $g''(\alpha) = 2(\alpha+1)^{-3} > 0$ so that $t = \alpha$ is the minimum of $g(t)$ with

$$g(\alpha) = \frac{\alpha}{1+\alpha} = \frac{\sigma^2}{\sigma^2 + a^2}.$$

It follows that

$$\Pr(X \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}.$$

Now, suppose that $\mu \neq 0$. Since $E(X - E(X)) = 0$ and $Var(X - E(X)) = Var(X) = \sigma^2$, by applying the previous inequality to $X - \mu$ we obtain

$$\Pr(X \geq \mu + a) \leq \frac{\sigma^2}{\sigma^2 + a^2}.$$

Similarly, since $E(\mu - X) = 0$ and $Var(\mu - X) = Var(X) = \sigma^2$, we get

$$\Pr(\mu - X \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

or

$$\Pr(X \leq \mu - a) \leq \frac{\sigma^2}{\sigma^2 + a^2} \blacksquare$$

Example 52.1
If the number produced in a factory during a week is a random variable with mean 100 and variance 400, compute an upper bound on the probability that this week's production will be at least 120.

Solution.
Applying the previous result we find

$$\Pr(X \geq 120) = \Pr(X - 100 \geq 20) \leq \frac{400}{400 + 20^2} = \frac{1}{2} \blacksquare$$

The following provides bounds on $\Pr(X \geq a)$ in terms of the moment generating function $M(t) = e^{tX}$ with $t > 0$.

Proposition 52.2 *(Chernoff's bound)*
Let X be a random variable and suppose that $M(t) = E(e^{tX})$ is finite. Then

$$\Pr(X \geq a) \leq e^{-ta}M(t), \quad t > 0$$

and

$$\Pr(X \leq a) \leq e^{-ta}M(t), \quad t < 0$$

Proof.
Suppose first that $t > 0$. Then

$$\begin{aligned}\Pr(X \geq a) \leq & \Pr(e^{tX} \geq e^{ta}) \\ \leq & E[e^{tX}]e^{-ta}\end{aligned}$$

where the last inequality follows from Markov's inequality. Similarly, for $t < 0$ we have

$$\begin{aligned}\Pr(X \leq a) \leq & \Pr(e^{tX} \geq e^{ta}) \\ \leq & E[e^{tX}]e^{-ta} \blacksquare\end{aligned}$$

It follows from Chernoff's inequality that a sharp bound for $\Pr(X \geq a)$ is a minimizer of the function $e^{-ta}M(t)$.

Example 52.2
Let Z is a standard random variable so that its moment generating function is $M(t) = e^{\frac{t^2}{2}}$. Find a sharp upper bound for $\Pr(Z \geq a)$.

Solution.
By Chernoff inequality we have

$$\Pr(Z \geq a) \leq e^{-ta}e^{\frac{t^2}{2}} = e^{\frac{t^2}{2}-ta}, \quad t > 0$$

Let $g(t) = e^{\frac{t^2}{2}-ta}$. Then $g'(t) = (t-a)e^{\frac{t^2}{2}-ta}$ so that $g'(t) = 0$ when $t = a$. Since $g''(t) = e^{\frac{t^2}{2}-ta} + (t-a)^2e^{\frac{t^2}{2}-ta}$ we find $g''(a) > 0$ so that $t = a$ is the minimum of $g(t)$. Hence, a sharp bound is

$$\Pr(Z \geq a) \leq e^{-\frac{a^2}{2}}, \quad t > 0$$

Similarly, for $a < 0$ we find

$$\Pr(Z \leq a) \leq e^{-\frac{a^2}{2}}, \quad t < 0 \blacksquare$$

The next inequality is one having to do with expectations rather than probabilities. Before stating it, we need the following definition: A differentiable function $f(x)$ is said to be **convex** on the open interval $I = (a, b)$ if

$$f(\alpha u + (1-\alpha)v) \leq \alpha f(u) + (1-\alpha)f(v)$$

for all u and v in I and $0 \leq \alpha \leq 1$. Geometrically, this says that the graph of $f(x)$ lies completely above each tangent line.

Proposition 52.3 *(Jensen's inequality)*
If $f(x)$ is a convex function then

$$E(f(X)) \geq f(E(X))$$

provided that the expectations exist and are finite.

Proof.
The tangent line at $(E(x), f(E(X)))$ is

$$y = f(E(X)) + f'(E(X))(x - E(X)).$$

By convexity we have

$$f(x) \geq f(E(X)) + f'(E(X))(x - E(X)).$$

Upon taking expectation of both sides we find

$$\begin{aligned} E(f(X)) \geq & E[f(E(X)) + f'(E(X))(X - E(X))] \\ = & f(E(X)) + f'(E(X))E(X) - f'(E(X))E(X) = f(E(X)) \blacksquare \end{aligned}$$

Example 52.3
Let X be a random variable. Show that $E(e^X) \geq e^{E(X)}$.

Solution.
Since $f(x) = e^x$ is convex, by Jensen's inequality we can write $E(e^X) \geq e^{E(X)}$ $\blacksquare$

Example 52.4
Suppose that $\{x_1, x_2, \cdots, x_n\}$ is a set of positive numbers. Show that the arithmetic mean is at least as large as the geometric mean:

$$(x_1 \cdot x_2 \cdots x_n)^{\frac{1}{n}} \leq \frac{1}{n}(x_1 + x_2 + \cdots + x_n).$$

Solution.
Let X be a random variable such that $\Pr(X = x_i) = \frac{1}{n}$ for $1 \leq i \leq n$. Let $g(x) = \ln x$. By Jensen's inequality we have

$$E[-\ln X] \geq -\ln[E(X)].$$

That is

$$E[\ln X] \leq \ln [E(X)].$$

But

$$E[\ln X] = \frac{1}{n}\sum_{i=1}^{n} \ln x_i = \frac{1}{n} \ln (x_1 \cdot x_2 \cdots x_n)$$

and

$$\ln [E(X)] = \ln \frac{1}{n}(x_1 + x_2 + \cdots + x_n).$$

It follows that

$$\ln (x_1 \cdot x_2 \cdots x_n)^{\frac{1}{n}} \leq \ln \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

Now the result follows by taking e^x of both sides and recalling that e^x is an increasing function ■

Practice Problems

Problem 52.1
Roll a single fair die and let X be the outcome. Then, $E(X) = 3.5$ and $Var(X) = \frac{35}{12}$.
(a) Compute the exact value of $\Pr(X \geq 6)$.
(b) Use Markov's inequality to find an upper bound of $\Pr(X \geq 6)$.
(c) Use Chebyshev's inequality to find an upper bound of $\Pr(X \geq 6)$.
(d) Use one-sided Chebyshev's inequality to find an upper bound of $\Pr(X \geq 6)$.

Problem 52.2
Find Chernoff bounds for a binomial random variable with parameters (n, p).

Problem 52.3
Suppose that the average number of sick kids in a pre-k class is three per day. Assume that the variance of the number of sick kids in the class in any one day is 9. Give an estimate of the probability that at least five kids will be sick tomorrow.

Problem 52.4
Suppose that you record only the integer amount of dollars of the checks you write in your checkbook. If 20 checks are written, what is an upper bound on the probability that the record in your checkbook shows at least \$15 less than the actual amount in your account?

Problem 52.5
Find the chernoff bounds for a Poisson random variable X with parameter λ.

Problem 52.6
Let X be a Poisson random variable with mean 20.
(a) Use the Markov's inequality to obtain an upper bound on

$$p = \Pr(X \geq 26).$$

(b) Use the Chernoff bound to obtain an upper bound on p.
(c) Use the Chebyshev's bound to obtain an upper bound on p.
(d) Approximate p by making use of the central limit theorem.

Problem 52.7
Let X be a random variable. Show the following
(a) $E(X^2) \geq [E(X)]^2$.
(b) If $X \geq 0$ then $E\left(\frac{1}{X}\right) \geq \frac{1}{E(X)}$.
(c) If $X > 0$ then $-E[\ln X] \geq -\ln[E(X)]$

Problem 52.8
Let X be a random variable with density function $f(x) = \frac{a}{x^{a+1}}, x \geq 1,\ a > 1$. We call X a **pareto** random variable with parameter a.
(a) Find $E(X)$.
(b) Find $E\left(\frac{1}{X}\right)$.
(c) Show that $g(x) = \frac{1}{x}$ is convex in $(0, \infty)$.
(d) Verify Jensen's inequality by comparing (b) and the reciprocal of (a).

Problem 52.9
Suppose that $\{x_1, x_2, \cdots, x_n\}$ is a set of positive numbers. Prove

$$(x_1 \cdot x_2 \cdots x_n)^{\frac{2}{n}} \leq \frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}.$$

Risk Management and Insurance

This section repersents a discussion of the study notes entitled "Risk and Insurance" by Anderson and Brown as listed in the SOA syllabus for Exam P.

By Economic risk or simply "risk" we mean one's possibility of losing economic security. For example, a driver faces a potential economic loss if his car is damaged and even a larger possible economic risk exists with respect to potential damages a driver might have to pay if he injures a third party in a car accident for which he is responsible.

Insurance is a form of risk management primarily used to hedge against the risk of a contingent, uncertain loss. Insurance is defined as the equitable transfer of the risk of a loss, from one entity (the insured) to another (the insurer), in exchange for payment. An **insurer** is a company selling the insurance; an **insured** or **policyholder** is the person or entity buying the insurance policy. The amount of money to be charged by the insurer for a certain amount of insurance coverage is called the **premium**. The insurance involves the insured assuming a guaranteed and known covered loss in the form of payment from the insurer upon the occurrence of a specific loss. The payment is referred to as the **benefit** or **claim payment**. This defined claim payment amount can be a fixed amount or can reimburse all or a part of the loss that occurred. The insured receives a contract called the **insurance policy** which details the conditions and circumstances under which the insured will be compensated.

Normally, only a small percentage of policyholders suffer losses. Their losses are paid out of the premiums collected from the pool of policyholders. Thus, the entire pool compensates the unfortunate few.

The Overall Loss Distribution Let X denote the overall loss of a policy. Let X_1 be the number of losses that will occur in a specified period. This random variable for the number of losses is commonly referred to as the **frequency of loss** and its probability distribution is called the **frequency distribution**. Let X_2 denote the amount of the loss, given that a loss has occurred. This random variable is often referred to as the **severity** and the probability distribution for the amount of loss is called the **severity distribution**.

Example 53.1
Consider a car owner who has an 80% chance of no accidents in a year, a 20% chance of being in a single accident in a year, and 0% chance of being in more than one accident. If there is an accident the severity distribution is given by the following table

X_2	Probability
500	0.50
5000	0.40
15000	0.10

(a) Calculate the total loss distribution function.
(b) Calculate the car owner's expected loss.
(c) Calculate the standard deviation of the annual loss incurred by the car owner.

Solution.
(a) Combining the frequency and severity distributions forms the following distribution of the random variable X, loss due to accident:

$$f(x) = \begin{cases} 0.80 & x = 0 \\ 20\%(0.50) = 0.10 & x = 500 \\ 20\%(0.40) = 0.08 & x = 5000 \\ 20\%(0.10) = 0.02 & x = 15000 \end{cases}$$

(b) The car owner's expected loss is

$$E(X) = 0.80 \times 0 + 0.10 \times 500 + 0.08 \times 5000 + 0.02 \times 15000 = \$750.$$

On average, the car owner spends 750 on repairs due to car accidents. A 750 loss may not seem like much to the car owner, but the possibility of a 5000

or 15,000 loss could create real concern.
(c) The standard deviation is

$$\begin{aligned}\sigma_X =&\sqrt{\sum(x-E(X))^2 f(x)}\\ =&\sqrt{0.80(-750)^2+0.10(-250)^2+0.08(4250)^2+0.02(14250)^2}\\ =&\sqrt{5962500}=2441.82 \blacksquare\end{aligned}$$

In all types of insurance there may be limits on benefits or claim payments. More specifically, there may be a maximum limit on the total reimbursed; there may be a minimum limit on losses that will be reimbursed; only a certain percentage of each loss may be reimbursed; or there may be different limits applied to particular types of losses. In each of these situations, the insurer does not reimburse the entire loss. Rather, the policyholder must cover part of the loss himself.
A policy may stipulate that losses are to be reimbursed only in excess of a stated threshold amount, called a **deductible.** For example, consider insurance that covers a loss resulting from an accident but includes a 500 deductible. If the loss is less than 500 the insurer will not pay anything to the policyholder. On the other hand, if the loss is more than 500, the insurer will pay for the loss in excess of the deductible. In other words, if the loss is 2000, the insurer will pay 1500.
Suppose that a insurance contract has a deductible of d and a maximum payment (i.e., benefit limit) of u per loss. Let X denote the total loss incurred by the policyholder and Y the payment received by the policyholder. Then

$$Y=\begin{cases} 0 & 0\le X\le d\\ X-d & d<X\le d+u\\ u & X\ge d+u\end{cases}$$

Example 53.2
Consider a policy with a deductible of 500 and benefit limit of 2500.
(a) How much will the policyholder receive if he/she suffered a loss of 450?
(b) How much will the policyholder receive if he/she suffered a loss of 1500?
(c) How much with the policyholder receive if he/she suffered a loss of 3500?

Solution.
(a) The loss is less than the deductible. As a result, the insurance company pays nothing and you need to cover the complete loss.

(b) You will cover 500 and the insurance company will pay you $1500-500=1000$.
(c) First, you have to cover 500. Since $3500-500=3000$ and the benefit limit is 2500, you will receive just 2500 from the insurance company. Thus, in total your share for the loss is 1000 and the insurance share is 2500 ■

Example 53.3
Consider the policy of Example 53.1. Suppose that the policy provides a 500 deductible and benefit limit of 12500. Calculte
(a) The annual expected payment made by the insurance company to a car owner.
(b) The standard deviation of the annual payment made by the insurance company to a car owner.
(c) The annual expected cost that the insured must cover out-of-pocket.
(d) The standard deviation of the annual expected cost that the insured must cover out-of-pocket.
(e) The correlation coefficient between insurer's annual payment and the insured's annual out-of-pocket cost to cover the loss.

Solution.
(a) Let Y denote the annual payment made by the insurance to the insurer. The pdf of this random variable is

$$f(y)=\begin{cases} 0.80 & y=0, x=0 \\ 20\%(0.50)=0.10 & y=0, x=500 \\ 20\%(0.40)=0.08 & y=4500, x=5000 \\ 20\%(0.10)=0.02 & y=12500, x=15000 \end{cases}$$

The annual expected payment made by the insurance to the insured is

$$E(Y)=0\times 0.90+4500\times 0.08+12500\times 0.02=610.$$

(b) The standard deviation of the annual payment made by the insurance company to a car owner is

$$\begin{aligned}\sigma_Y=&\sqrt{\sum(y-E(Y))^2f(y)}\\=&\sqrt{0.90(-610)^2+0.08(3890)^2+0.02(11890)^2}=\sqrt{4372900}=2091.15\end{aligned}$$

(c) Let Z be the annual expected cost that the insured must cover out-of-pocket.The distribution of Z is

$$f(z) = \begin{cases} 0.80 & z = 0, x = 0 \\ 20\%(0.50) = 0.10 & z = 500, x = 500 \\ 20\%(0.40) = 0.08 & z = 500, x = 5000 \\ 20\%(0.10) = 0.02 & z = 2500, x = 15000 \end{cases}$$

the annual expected cost that the insured must cover out-of-pocket is

$$E(Z) = 0.10(500) + 0.08(500) + 0.02(2500) = 140.$$

(d) The standard deviation of the annual expected cost that the insured must cover out-of-pocket is

$$\begin{aligned}\sigma_Z =& \sqrt{\sum (z - E(Z))^2 f(z)} \\ =& \sqrt{0.80(-140)^2 + 0.18(360)^2 + 0.02(2360)^2} = \sqrt{150400} = 387.81\end{aligned}$$

(e) We have that $X = Y + Z$ where X as defined in Example 53.1. From the formula

$$\text{Var}(X) = \text{Var}(Y) + \text{Var}(Z) - 2\text{Cov}(Y, Z)$$

we find

$$\text{Cov}(Y, Z) = \frac{5962500 - 4372900 - 150400}{2} = 719600.$$

Finally,

$$\rho_{Y,Z} = \frac{\text{Cov}(Y, Z)}{\sigma_Y \sigma_Z} = \frac{719600}{2091.15 \times 387.81} = 0.8873 \blacksquare$$

Continuous Severity Distributions

In the car insurance example, Example 53.1, the severity distribution was a discrete distribution. In what follows, we consider the continuous case of the severity distribution.

Example 53.4

Consider an insurance policy that reimburses annual hospital charges for an insured individual. The probability of any individual being hospitalized in a year is 15%. Once an individual is hospitalized, the charges X have a probability density function $f_X(x|H = 1) = 0.1e^{-0.1x}$ for $x > 0$ and $f_X(x) = 0$ for $x \leq 0$. Determine the expected value, the standard deviation, and the ratio of the standard deviation to the mean (coefficient of variation) of hospital charges for an insured individual.

Solution.
The expected value of hospital charges is

$$\begin{aligned}E(X) =&\Pr(H \neq 1)E[X|H \neq 1] + \Pr(H = 1)E[X|H = 1]\\ =&0.85 \times 0 + 0.15\int_0^\infty 0.1xe^{-0.1x}dx = 1.5\end{aligned}$$

Now,

$$\begin{aligned}E(X^2) =&\Pr(H \neq 1)E[X^2|H \neq 1] + \Pr(H = 1)E[X^2|H = 1]\\ =&0.85 \times 0^2 + 0.15\int_0^\infty 0.1x^2e^{-0.1x}dx = 30\end{aligned}$$

The variance of the hospital charges is given by

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = 30 - 1.5^2 = 27.75$$

so that the standard deviation is $\sigma_X = \sqrt{27.75} = 5.27$. Finally, the coefficient of variation is

$$\frac{\sigma_X}{E(X)} = \frac{5.27}{0.15} = 3.51 \blacksquare$$

Example 53.5
Using the previous example, determine the expected claim payments, standard deviation and coefficient of variation for an insurance pool that reimburses hospital charges for 200 individuals. Assume that claims for each individual are independent of the other individuals.

Solution.
Let $S = X_1 + X_2 + \cdots + X_{200}$. Since the claims are independent, we have

$$\begin{aligned}E(S) =&200E(X) = 200 \times 1.5 = 300\\ \sigma_S =&10\sqrt{2}\sigma_X = 74.50\end{aligned}$$

and the coefficient of variation is

$$\frac{\sigma_S}{E(S)} = \frac{74.50}{300} = 0.25 \blacksquare$$

Example 53.6
In Example 53.4, assume that there is a deductible of 5. Determine the expected value, standard deviation and coefficient of variation of the claim payment.

Solution.
Let Y represent claim payments to hospital charges. Letting $Z = \max(0, X - 5)$ we can write

$$Y = \begin{cases} Z & \text{with probability } 0.15 \\ 0 & \text{with probability } 0.85 \end{cases}$$

The expected value of the claim payments to hospital charges is

$$\begin{aligned} E(Y) =& 0.85 \times 0 + 0.15 \times E(Z) \\ =& 0.15 \int_0^\infty \max(0, x-5) f_X(x) dx \\ =& 0.15 \int_5^\infty 0.1(x-5)e^{-0.1x} dx \\ =& 1.5e^{-0.5} \end{aligned}$$

Likewise,

$$\begin{aligned} E(Y^2) =& 0.85 \times 0^2 + 0.15 \times E(Z^2) \\ =& 0.15 \int_0^\infty \max(0, x-5)^2 f_X(x) dx \\ =& 0.15 \int_5^\infty 0.1(x-5)^2 e^{-0.1x} dx \\ =& 30e^{-0.5} \end{aligned}$$

The variance of the claim payments to hospital charges is

$$\text{Var}(Y) = E(Y^2) - [E(Y)]^2 = 30e^{-0.5} - 2.25e^{-1} = 17.3682$$

and the standard deviation is $\sigma_Y = \sqrt{17.3682} = 4.17$. Finally, the coefficient of variation is

$$\frac{\sigma_Y}{E(Y)} = \frac{4.17}{1.5e^{-0.5}} = 4.58 \blacksquare$$

Practice Problems

Problem 53.1

Consider a policy with a deductible of 200 and benefit limit of 5000. The policy states that the insurer will pay 90% of the loss in excess of the deductible subject to the benefit claim
(a) How much will the policyholder receive if he/she suffered a loss of 4000?
(b) How much will the policyholder receive if he/she suffered a loss of 5750?
(c) How much with the policyholder receive if he/she suffered a loss of 5780?

Problem 53.2

Consider a car owner who has an 80% chance of no accidents in a year, a 20% chance of being in a single accident in a year, and 0% chance of being in more than one accident. If there is an accident the severity distribution is given by the following table

X_2	Probability
500	0.50
5000	0.40
15000	0.10

There is an annual deductible of 500 and the annual maximum payment by the insurer is 12500. The insurer will pay 40% of the loss in excess of the deductible subject to the maximum annual payment. Calculate
(a) The distribution function of the random variable Y representing the payment made by the insurer to the insured.
(b) The annual expected payment made by the insurance company to a car owner.
(c) The standard deviation of the annual payment made by the insurance company to a car owner.
(d) The annual expected cost that the insured must cover out-of-pocket.
(e) The standard deviation of the annual expected cost that the insured must cover out-of-pocket.
(f) The correlation coefficient between insurer's annual payment and the insured's annual out-of-pocket cost to cover the loss.

Problem 53.3 ‡

Automobile losses reported to an insurance company are independent and uniformly distributed between 0 and 20,000. The company covers each such

loss subject to a deductible of 5,000.
Calculate the probability that the total payout on 200 reported losses is between 1,000,000 and 1,200,000.

Problem 53.4 ‡
The amount of a claim that a car insurance company pays out follows an exponential distribution. By imposing a deductible of d, the insurance company reduces the expected claim payment by 10%.
Calculate the percentage reduction on the variance of the claim payment.

Sample Exam 1

Problem 1 ‡
A survey of a group's viewing habits over the last year revealed the following information

(i)	28% watched gymnastics
(ii)	29% watched baseball
(iii)	19% watched soccer
(iv)	14% watched gymnastics and baseball
(v)	12% watched baseball and soccer
(vi)	10% watched gymnastics and soccer
(vii)	8% watched all three sports.

Calculate the percentage of the group that watched none of the three sports during the last year.
(A) 24
(B) 36
(C) 41
(D) 52
(E) 60

Problem 2 ‡
An insurance company estimates that 40% of policyholders who have only an auto policy will renew next year and 60% of policyholders who have only a homeowners policy will renew next year. The company estimates that 80% of policyholders who have both an auto and a homeowners policy will renew at least one of those policies next year. Company records show that 65% of policyholders have an auto policy, 50% of policyholders have a homeowners policy, and 15% of policyholders have both an auto and a homeowners policy. Using the company's estimates, calculate the percentage of policyholders that will renew at least one policy next year.

(A) 20
(B) 29
(C) 41
(D) 53
(E) 70

Problem 3 ‡
An insurer offers a health plan to the employees of a large company. As part of this plan, the individual employees may choose exactly two of the supplementary coverages A, B, and C, or they may choose no supplementary coverage. The proportions of the company's employees that choose coverages A, B, and C are $\frac{1}{4}, \frac{1}{3}$, and , $\frac{5}{12}$ respectively.
Determine the probability that a randomly chosen employee will choose no supplementary coverage.

(A) 0
(B) $\frac{47}{144}$
(C) $\frac{1}{2}$
(D) $\frac{97}{144}$
(E) $\frac{7}{9}$

Problem 4 ‡
An insurance agent offers his clients auto insurance, homeowners insurance and renters insurance. The purchase of homeowners insurance and the purchase of renters insurance are mutually exclusive. The profile of the agent's clients is as follows:
i) 17% of the clients have none of these three products.
ii) 64% of the clients have auto insurance.
iii) Twice as many of the clients have homeowners insurance as have renters insurance.
iv) 35% of the clients have two of these three products.
v) 11% of the clients have homeowners insurance, but not auto insurance.
Calculate the percentage of the agent's clients that have both auto and renters insurance.

(A) 7%
(B) 10%

(C) 16%
(D) 25%
(E) 28

Problem 5 ‡
From 27 pieces of luggage, an airline luggage handler damages a random sample of four. The probability that exactly one of the damaged pieces of luggage is insured is twice the probability that none of the damaged pieces are insured. Calculate the probability that exactly two of the four damaged pieces are insured.

(A) 0.06
(B) 0.13
(C) 0.27
(D) 0.30
(E) 0.31

Problem 6 ‡
An auto insurance company insures drivers of all ages. An actuary compiled the following statistics on the company's insured drivers:

Age of Driver	Probability of Accident	Portion of Company's Insured Drivers
16 - 20	0.06	0.08
21 - 30	0.03	0.15
31 - 65	0.02	0.49
66 - 99	0.04	0.28

A randomly selected driver that the company insures has an accident. Calculate the probability that the driver was age 16-20.

(A) 0.13
(B) 0.16
(C) 0.19
(D) 0.23
(E) 0.40

Problem 7 ‡
An actuary studied the likelihood that different types of drivers would be

involved in at least one collision during any one-year period. The results of the study are presented below.

Type of driver	Percentage of all drivers	Probability of at least one collision
Teen	8%	0.15
Young adult	16%	0.08
Midlife	45%	0.04
Senior	31%	0.05
Total	100%	

Given that a driver has been involved in at least one collision in the past year, what is the probability that the driver is a young adult driver?

(A) 0.06
(B) 0.16
(C) 0.19
(D) 0.22
(E) 0.25

Problem 8 ‡
Ten percent of a company's life insurance policyholders are smokers. The rest are nonsmokers. For each nonsmoker, the probability of dying during the year is 0.01. For each smoker, the probability of dying during the year is 0.05.
Given that a policyholder has died, what is the probability that the policyholder was a smoker?

(A) 0.05
(B) 0.20
(C) 0.36
(D) 0.56
(E) 0.90

Problem 9 ‡
Workplace accidents are categorized into three groups: minor, moderate and severe. The probability that a given accident is minor is 0.5, that it is moderate is 0.4, and that it is severe is 0.1. Two accidents occur independently

in one month.
Calculate the probability that neither accident is severe and at most one is moderate.

(A) 0.25
(B) 0.40
(C) 0.45
(D) 0.56
(E) 0.65

Problem 10 ‡
Two life insurance policies, each with a death benefit of 10,000 and a one-time premium of 500, are sold to a couple, one for each person. The policies will expire at the end of the tenth year. The probability that only the wife will survive at least ten years is 0.025, the probability that only the husband will survive at least ten years is 0.01, and the probability that both of them will survive at least ten years is 0.96 .
What is the expected excess of premiums over claims, given that the husband survives at least ten years?

(A) 350
(B) 385
(C) 397
(D) 870
(E) 897

Problem 11 ‡
A probability distribution of the claim sizes for an auto insurance policy is given in the table below:

Claim size	Probability
20	0.15
30	0.10
40	0.05
50	0.20
60	0.10
70	0.10
80	0.30

What percentage of the claims are within one standard deviation of the mean claim size?

(A) 45%
(B) 55%
(C) 68%
(D) 85%
(E) 100%

Problem 12 ‡
A company prices its hurricane insurance using the following assumptions:

(i) In any calendar year, there can be at most one hurricane.
(ii) In any calendar year, the probability of a hurricane is 0.05 .
(iii) The number of hurricanes in any calendar year is independent of the number of hurricanes in any other calendar year.

Using the company's assumptions, calculate the probability that there are fewer than 3 hurricanes in a 20-year period.

(A) 0.06
(B) 0.19
(C) 0.38
(D) 0.62
(E) 0.92

Problem 13 ‡
A company buys a policy to insure its revenue in the event of major snowstorms that shut down business. The policy pays nothing for the first such snowstorm of the year and $10,000 for each one thereafter, until the end of the year. The number of major snowstorms per year that shut down business is assumed to have a Poisson distribution with mean 1.5 .
What is the expected amount paid to the company under this policy during a one-year period?

(A) 2,769
(B) 5,000
(C) 7,231

(D) 8,347
(E) 10,578

Problem 14 ‡
Each time a hurricane arrives, a new home has a 0.4 probability of experiencing damage. The occurrences of damage in different hurricanes are independent. Calculate the mode of the number of hurricanes it takes for the home to experience damage from two hurricanes. Hint: The mode of X is the number that maximizes the probability mass function of X.

(A) 2
(B) 3
(C) 4
(D) 5
(E) 6

Problem 15 ‡
An insurance company insures a large number of homes. The insured value, X, of a randomly selected home is assumed to follow a distribution with density function

$$f(x) = \begin{cases} 3x^{-4} & x > 1 \\ 0 & \text{otherwise} \end{cases}$$

Given that a randomly selected home is insured for at least 1.5, what is the probability that it is insured for less than 2 ?

(A) 0.578
(B) 0.684
(C) 0.704
(D) 0.829
(E) 0.875

Problem 16 ‡
A manufacturer's annual losses follow a distribution with density function

$$f(x) = \begin{cases} \frac{2.5(0.6)^{2.5}}{x^{3.5}} & x > 0.6 \\ 0 & \text{otherwise.} \end{cases}$$

To cover its losses, the manufacturer purchases an insurance policy with an annual deductible of 2.

What is the mean of the manufacturer's annual losses not paid by the insurance policy?

(A) 0.84
(B) 0.88
(C) 0.93
(D) 0.95
(E) 1.00

Problem 17 ‡
A random variable X has the cumulative distribution function

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{x^2-2x+2}{2} & 1 \le x < 2 \\ 1 & x \ge 2. \end{cases}$$

Calculate the variance of X.

(A) $\frac{7}{72}$
(B) $\frac{1}{8}$
(C) $\frac{5}{36}$
(D) $\frac{4}{3}$
(E) $\frac{23}{12}$

Problem 18 ‡
An insurance company sells an auto insurance policy that covers losses incurred by a policyholder, subject to a deductible of 100 . Losses incurred follow an exponential distribution with mean 300.
What is the 95th percentile of actual losses that exceed the deductible?

(A) 600
(B) 700
(C) 800
(D) 900
(E) 1000

Problem 19 ‡
The owner of an automobile insures it against damage by purchasing an insurance policy with a deductible of 250 . In the event that the automobile is

damaged, repair costs can be modeled by a uniform random variable on the interval $(0, 1500)$.
Determine the standard deviation of the insurance payment in the event that the automobile is damaged.

(A) 361
(B) 403
(C) 433
(D) 464
(E) 521

Problem 20 ‡
The lifetime of a printer costing 200 is exponentially distributed with mean 2 years. The manufacturer agrees to pay a full refund to a buyer if the printer fails during the first year following its purchase, and a one-half refund if it fails during the second year.
If the manufacturer sells 100 printers, how much should it expect to pay in refunds?

(A) 6,321
(B) 7,358
(C) 7,869
(D) 10,256
(E) 12,642

Problem 21 ‡
The time to failure of a component in an electronic device has an exponential distribution with a median of four hours.
Calculate the probability that the component will work without failing for at least five hours.

(A) 0.07
(B) 0.29
(C) 0.38
(D) 0.42
(E) 0.57

Problem 22 ‡
An actuary models the lifetime of a device using the random variable $Y =$

$10X^{0.8}$, where X is an exponential random variable with mean 1 year. Determine the probability density function $f_Y(y)$, for $y > 0$, of the random variable Y.

(A) $10y^{0.8}e^{-8y^{-0.2}}$
(B) $8y^{-0.2}e^{-10y^{0.8}}$
(C) $8y^{-0.2}e^{-(0.1y)^{1.25}}$
(D) $(0.1y)^{1.25}e^{-0.125(0.1y)^{0.25}}$
(E) $0.125(0.1y)^{0.25}e^{-(0.1y)^{1.25}}$

Problem 23 ‡
A device runs until either of two components fails, at which point the device stops running. The joint density function of the lifetimes of the two components, both measured in hours, is

$$f_{XY}(x,y) = \begin{cases} \frac{x+y}{8} & 0 < x, y < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the probability that the device fails during its first hour of operation?

(A) 0.125
(B) 0.141
(C) 0.391
(D) 0.625
(E) 0.875

Problem 24 ‡
Let X represent the age of an insured automobile involved in an accident. Let Y represent the length of time the owner has insured the automobile at the time of the accident. X and Y have joint probability density function

$$f_{XY}(x,y) = \begin{cases} \frac{1}{64}(10 - xy^2) & 2 \le x \le 10, 0 \le y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the expected age of an insured automobile involved in an accident.

(A) 4.9
(B) 5.2
(C) 5.8
(D) 6.0
(E) 6.4

Problem 25 ‡
A device contains two components. The device fails if either component fails. The joint density function of the lifetimes of the components, measured in hours, is $f(s,t)$, where $0 < s < 1$ and $0 < t < 1$.
What is the probability that the device fails during the first half hour of operation?

(A) $\int_0^{0.5}\int_0^{0.5} f(s,t)dsdt$

(B) $\int_0^1\int_0^{0.5} f(s,t)dsdt$

(C) $\int_{0.5}^1\int_{0.5}^1 f(s,t)dsdt$

(D) $\int_0^{0.5}\int_0^1 f(s,t)dsdt + \int_0^1\int_0^{0.5} f(s,t)dsdt$

(E) $\int_0^{0.5}\int_{0.5}^1 f(s,t)dsdt + \int_0^1\int_0^{0.5} f(s,t)dsdt$

Problem 26 ‡
The waiting time for the first claim from a good driver and the waiting time for the first claim from a bad driver are independent and follow exponential distributions with means 6 years and 3 years, respectively.
What is the probability that the first claim from a good driver will be filed within 3 years and the first claim from a bad driver will be filed within 2 years?

(A) $\frac{1}{18}(1 - e^{-\frac{2}{3}} - e^{-\frac{1}{2}} + e^{-\frac{7}{6}})$
(B) $\frac{1}{18}e^{-\frac{7}{6}}$
(C) $1 - e^{-\frac{2}{3}} - e^{-\frac{1}{2}} + e^{-\frac{7}{6}}$
(D) $1 - e^{-\frac{2}{3}} - e^{-\frac{1}{2}} + e^{-\frac{1}{3}}$
(E) $1 - \frac{1}{3}e^{-\frac{2}{3}} - \frac{1}{6}e^{-\frac{1}{2}} + \frac{1}{18}e^{-\frac{7}{6}}$

Problem 27 ‡
In a small metropolitan area, annual losses due to storm, fire, and theft are assumed to be independent, exponentially distributed random variables with respective means 1.0, 1.5, and 2.4 .
Determine the probability that the maximum of these losses exceeds 3.

(A) 0.002

(B) 0.050
(C) 0.159
(D) 0.287
(E) 0.414

Problem 28 ‡
Let N_1 and N_2 represent the numbers of claims submitted to a life insurance company in April and May, respectively. The joint probability function of N_1 and N_2 is

$$Pr(n_1, n_2) = \begin{cases} \frac{3}{4}\left(\frac{1}{4}\right)^{n_1-1} e^{-n_1}(1-e^{-n_1})^{n_2-1}, & \text{for } n_1 = 1,2,3,\cdots \text{ and } n_2 = 1,2,3,\cdots \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the expected number of claims that will be submitted to the company in May if exactly 2 claims were submitted in April.

(A) $\frac{3}{16}(e^2-1)$
(B) $\frac{3}{16}e^2$
(C) $\frac{3e}{4-e}$
(D) e^2-1
(E) e^2

Problem 29 ‡
An auto insurance policy will pay for damage to both the policyholder's car and the other driver's car in the event that the policyholder is responsible for an accident. The size of the payment for damage to the policyholder's car, X, has a marginal density function of 1 for $0 < x < 1$. Given $X = x$, the size of the payment for damage to the other driver's car, Y, has conditional density of 1 for $x < y < x + 1$.
If the policyholder is responsible for an accident, what is the probability that the payment for damage to the other driver's car will be greater than 0.5?

(A) $\frac{3}{8}$
(B) $\frac{1}{2}$
(C) $\frac{3}{4}$
(D) $\frac{7}{8}$
(E) $\frac{15}{16}$

Problem 30 ‡
Let T_1 be the time between a car accident and reporting a claim to the insurance company. Let T_2 be the time between the report of the claim and payment of the claim. The joint density function of T_1 and T_2, $f(t_1, t_2)$, is constant over the region $0 < t_1 < 6, 0 < t_2 < 6, t_1 + t_2 < 10$, and zero otherwise.
Determine $E[T_1 + T_2]$, the expected time between a car accident and payment of the claim.

(A) 4.9
(B) 5.0
(C) 5.7
(D) 6.0
(E) 6.7

Problem 31 ‡
An insurance policy pays a total medical benefit consisting of two parts for each claim. Let X represent the part of the benefit that is paid to the surgeon, and let Y represent the part that is paid to the hospital. The variance of X is 5000, the variance of Y is 10,000, and the variance of the total benefit, $X + Y$, is 17,000.
Due to increasing medical costs, the company that issues the policy decides to increase X by a flat amount of 100 per claim and to increase Y by 10% per claim.
Calculate the variance of the total benefit after these revisions have been made.

(A) 18,200
(B) 18,800
(C) 19,300
(D) 19,520
(E) 20,670

Problem 32 ‡
Let X and Y be continuous random variables with joint density function

$$f_{XY}(x, y) = \begin{cases} \frac{8}{3}xy & 0 \le x \le 1, x \le y \le 2x \\ 0 & \text{otherwise} \end{cases}$$

Find Cov(X, Y)

(A) 0.04
(B) 0.25
(C) 0.67
(D) 0.80
(E) 1.24

Problem 33 ‡
A diagnostic test for the presence of a disease has two possible outcomes: 1 for disease present and 0 for disease not present. Let X denote the disease state of a patient, and let Y denote the outcome of the diagnostic test. The joint probability function of X and Y is given by:

$$\begin{aligned}Pr(X=0,Y=0)=&0.800\\Pr(X=1,Y=0)=&0.050\\Pr(X=0,Y=1)=&0.025\\Pr(X=1,Y=1)=&0.125.\end{aligned}$$

Calculate $Var(Y|X=1)$.

(A) 0.13
(B) 0.15
(C) 0.20
(D) 0.51
(E) 0.71

Problem 34 ‡
The number of workplace injuries, N, occurring in a factory on any given day is Poisson distributed with mean λ. The parameter λ is a random variable that is determined by the level of activity in the factory, and is uniformly distributed on the interval $[0, 3]$.
Calculate Var(N).

(A) λ
(B) 2λ
(C) 0.75
(D) 1.50
(E) 2.25

Problem 35 ‡
A motorist makes three driving errors, each independently resulting in an accident with probability 0.25.
Each accident results in a loss that is exponentially distributed with mean 0.80. Losses are mutually independent and independent of the number of accidents. The motorist's insurer reimburses 70% of each loss due to an accident.
Calculate the variance of the total unreimbursed loss the motorist experiences due to accidents resulting from these driving errors.

(A) 0.0432
(B) 0.0756
(C) 0.1782
(D) 0.2520
(E) 0.4116

Problem 36 ‡
An actuary determines that the claim size for a certain class of accidents is a random variable, X, with moment generating function

$$M_X(t) = \frac{1}{(1-2500t)^4}$$

Determine the standard deviation of the claim size for this class of accidents.

(A) 1,340
(B) 5,000
(C) 8,660
(D) 10,000
(E) 11,180

Problem 37 ‡
The value of a piece of factory equipment after three years of use is $100(0.5)^X$ where X is a random variable having moment generating function

$$M_X(t) = \tfrac{1}{1-2t} \text{ for } t < \tfrac{1}{2}.$$

Calculate the expected value of this piece of equipment after three years of use.

(A) 12.5
(B) 25.0
(C) 41.9
(D) 70.7
(E) 83.8

Problem 38 ‡
An insurance company issues 1250 vision care insurance policies. The number of claims filed by a policyholder under a vision care insurance policy during one year is a Poisson random variable with mean 2. Assume the numbers of claims filed by distinct policyholders are independent of one another.
What is the approximate probability that there is a total of between 2450 and 2600 claims during a one-year period?

(A) 0.68
(B) 0.82
(C) 0.87
(D) 0.95
(E) 1.00

Problem 39 ‡
A city has just added 100 new female recruits to its police force. The city will provide a pension to each new hire who remains with the force until retirement. In addition, if the new hire is married at the time of her retirement, a second pension will be provided for her husband. A consulting actuary makes the following assumptions:

(i) Each new recruit has a 0.4 probability of remaining with the police force until retirement.
(ii) Given that a new recruit reaches retirement with the police force, the probability that she is not married at the time of retirement is 0.25.
(iii) The number of pensions that the city will provide on behalf of each new hire is independent of the number of pensions it will provide on behalf of any other new hire.
Determine the probability that the city will provide at most 90 pensions to the 100 new hires and their husbands.

(A) 0.60

(B) 0.67
(C) 0.75
(D) 0.93
(E) 0.99

Answers

1. D
2. D
3. C
4. B
5. C
6. B
7. D
8. C
9. E
10. E
11. A
12. E
13. C
14. B
15. A
16. C
17. C
18. D
19. B
20. D
21. D
22. E
23. D
24. C
25. E
26. C
27. E
28. E
29. D
30. C
31. C
32. A
33. C
34. E
35. B
36. B

37. C
38. B
39. E

Sample Exam 2

Problem 1 ‡
An auto insurance has 10,000 policyholders. Each policyholder is classified as

(i) young or old;
(ii) male or female;
(iii) married or single.

Of these policyholders, 3,000 are young, 4,600 are male, and 7,000 are married. The policyholders can also be classified as 1,320 young males, 3,010 married males, and 1,400 young married persons. Finally, 600 of the policyholders are young married males.
How many of the company's policyholders are young, female, and single?

(A) 280
(B) 423
(C) 486
(D) 880
(E) 896

Problem 2 ‡
A doctor is studying the relationship between blood pressure and heartbeat abnormalities in her patients. She tests a random sample of her patients and notes their blood pressures (high, low, or normal) and their heartbeats (regular or irregular). She finds that:

(i) 14% have high blood pressure.
(ii) 22% have low blood pressure.
(iii) 15% have an irregular heartbeat.
(iv) Of those with an irregular heartbeat, one-third have high blood pressure.
(v) Of those with normal blood pressure, one-eighth have an irregular heartbeat.

What portion of the patients selected have a regular heartbeat and low blood pressure?

(A) 2%
(B) 5%
(C) 8%
(D) 9%
(E) 20%

Problem 3 ‡
You are given $\Pr(A \cup B) = 0.7$ and $\Pr(A \cup B^c) = 0.9$. Determine $\Pr(A)$.

(A) 0.2
(B) 0.3
(C) 0.4
(D) 0.6
(E) 0.8

Problem 4 ‡
A mattress store sells only king, queen and twin-size mattresses. Sales records at the store indicate that one-fourth as many queen-size mattresses are sold as king and twin-size mattresses combined. Records also indicate that three times as many king-size mattresses are sold as twin-size mattresses.
Calculate the probability that the next mattress sold is either king or queen-size.

(A) 0.12
(B) 0.15
(C) 0.80
(D) 0.85
(E) 0.95

Problem 5 ‡
A public health researcher examines the medical records of a group of 937 men who died in 1999 and discovers that 210 of the men died from causes related to heart disease. Moreover, 312 of the 937 men had at least one parent who suffered from heart disease, and, of these 312 men, 102 died from causes related to heart disease.
Determine the probability that a man randomly selected from this group died of causes related to heart disease, given that neither of his parents suffered from heart disease.

(A) 0.115
(B) 0.173
(C) 0.224
(D) 0.327
(E) 0.514

Problem 6 ‡
An insurance company issues life insurance policies in three separate categories: standard, preferred, and ultra-preferred. Of the company's policyholders, 50% are standard, 40% are preferred, and 10% are ultra-preferred. Each standard policyholder has probability 0.010 of dying in the next year, each preferred policyholder has probability 0.005 of dying in the next year, and each ultra-preferred policyholder has probability 0.001 of dying in the next year.
A policyholder dies in the next year. What is the probability that the deceased policyholder was ultra-preferred?

(A) 0.0001
(B) 0.0010
(C) 0.0071
(D) 0.0141
(E) 0.2817

Problem 7 ‡
A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence of the disease 0.5% of the time when the disease is not present. One percent of the population actually has the disease.

Calculate the probability that a person has the disease given that the test indicates the presence of the disease.

(A) 0.324
(B) 0.657
(C) 0.945
(D) 0.950
(E) 0.995

Problem 8 ‡
One urn contains 4 red balls and 6 blue balls. A second urn contains 16 red balls and x blue balls. A single ball is drawn from each urn. The probability that both balls are the same color is 0.44 .
Calculate x.

(A) 4
(B) 20
(C) 24
(D) 44
(E) 64

Problem 9 ‡
The number of injury claims per month is modeled by a random variable N with

$$\Pr(N = n) = \frac{1}{(n+1)(n+2)}, \quad n \geq 0.$$

Determine the probability of at least one claim during a particular month, given that there have been at most four claims during that month.

(A) $\frac{1}{3}$
(B) $\frac{2}{5}$
(C) $\frac{1}{2}$
(D) $\frac{3}{5}$
(E) $\frac{5}{6}$

Problem 10 ‡
An auto insurance company is implementing a new bonus system. In each month, if a policyholder does not have an accident, he or she will receive a

5.00 cash-back bonus from the insurer.
Among the 1,000 policyholders of the auto insurance company, 400 are classified as low-risk drivers and 600 are classified as high-risk drivers.
In each month, the probability of zero accidents for high-risk drivers is 0.80 and the probability of zero accidents for low-risk drivers is 0.90.
Calculate the expected bonus payment from the insurer to the 1000 policyholders in one year.

(A) 48,000
(B) 50,400
(C) 51,000
(D) 54,000
(E) 60,000

Problem 11 ‡
The annual cost of maintaining and repairing a car averages 200 with a variance of 260. what will be the variance of the annual cost of maintaining and repairing a car if 20% tax is introduced on all items associated with the maintenance and repair of cars?

(A) 208
(B) 260
(C) 270
(D) 312
(E) 374

Problem 12 ‡
A tour operator has a bus that can accommodate 20 tourists. The operator knows that tourists may not show up, so he sells 21 tickets. The probability that an individual tourist will not show up is 0.02, independent of all other tourists.
Each ticket costs 50, and is non-refundable if a tourist fails to show up. If a tourist shows up and a seat is not available, the tour operator has to pay 100 (ticket cost + 50 penalty) to the tourist.
What is the expected revenue of the tour operator?

(A) 935
(B) 950

(C) 967
(D) 976
(E) 985

Problem 13 ‡
A baseball team has scheduled its opening game for April 1. If it rains on April 1, the game is postponed and will be played on the next day that it does not rain. The team purchases insurance against rain. The policy will pay 1000 for each day, up to 2 days, that the opening game is postponed.
The insurance company determines that the number of consecutive days of rain beginning on April 1 is a Poisson random variable with mean 0.6 .
What is the standard deviation of the amount the insurance company will have to pay?

(A) 668
(B) 699
(C) 775
(D) 817
(E) 904

Problem 14 ‡
The lifetime of a machine part has a continuous distribution on the interval $(0, 40)$ with probability density function f, where $f(x)$ is proportional to $(10 + x)^{-2}$.
Calculate the probability that the lifetime of the machine part is less than 6.

(A) 0.04
(B) 0.15
(C) 0.47
(D) 0.53
(E) 0.94

Problem 15 ‡
An insurance policy pays for a random loss X subject to a deductible of C, where $0 < C < 1$. The loss amount is modeled as a continuous random variable with density function

$$f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Given a random loss X, the probability that the insurance payment is less than 0.5 is equal to 0.64 . Calculate C.

(A) 0.1
(B) 0.3
(C) 0.4
(D) 0.6
(E) 0.8

Problem 16 ‡
Let X be a continuous random variable with density function

$$f(x) = \begin{cases} \frac{|x|}{10} & -2 \leq x \leq 4 \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the expected value of X.

(A) $\frac{1}{5}$
(B) $\frac{3}{5}$
(C) 1
(D) $\frac{28}{15}$
(E) $\frac{12}{5}$

Problem 17 ‡
A company agrees to accept the highest of four sealed bids on a property. The four bids are regarded as four independent random variables with common cumulative distribution function

$$F(x) = \frac{1}{2}(1 + \sin \pi x), \quad \frac{3}{2} \leq x \leq \frac{5}{2}$$

and 0 otherwise. What is the expected value of the accepted bid?

(A) $\pi \int_{\frac{3}{2}}^{\frac{5}{2}} x \cos(\pi x) dx$
(B) $\frac{1}{16} \int_{\frac{3}{2}}^{\frac{5}{2}} (1 + \sin(\pi x))^4 dx$
(C) $\frac{1}{16} \int_{\frac{3}{2}}^{\frac{5}{2}} x(1 + \sin(\pi x))^4 dx$
(D) $\frac{1}{4}\pi \int_{\frac{3}{2}}^{\frac{5}{2}} \cos(\pi x)(1 + \sin(\pi x))^3 dx$
(E) $\frac{1}{4}\pi \int_{\frac{3}{2}}^{\frac{5}{2}} x \cos(\pi x)(1 + \sin(\pi x))^3 dx$

Problem 18 ‡
An automobile insurance company issues a one-year policy with a deductible of 500. The probability is 0.8 that the insured automobile has no accident and 0.0 that the automobile has more than one accident. If there is an accident, the loss before application of the deductible is exponentially distributed with mean 3000.
Calculate the 95$^{\text{th}}$ percentile of the insurance company payout on this policy.

(A) 3466
(B) 3659
(C) 4159
(D) 8487
(E) 8987

Problem 19 ‡
For Company A there is a 60% chance that no claim is made during the coming year. If one or more claims are made, the total claim amount is normally distributed with mean 10,000 and standard deviation 2,000 .
For Company B there is a 70% chance that no claim is made during the coming year. If one or more claims are made, the total claim amount is normally distributed with mean 9,000 and standard deviation 2,000 .
Assuming that the total claim amounts of the two companies are independent, what is the probability that, in the coming year, Company B's total claim amount will exceed Company A's total claim amount?

(A) 0.180
(B) 0.185
(C) 0.217
(D) 0.223
(E) 0.240

Problem 20 ‡
A device that continuously measures and records seismic activity is placed in a remote region. The time, T, to failure of this device is exponentially distributed with mean 3 years. Since the device will not be monitored during its first two years of service, the time to discovery of its failure is $X = \max(T, 2)$.
Determine $E[X]$.

(A) $2+\frac{1}{3}e^{-6}$
(B) $2-2e^{-\frac{2}{3}}+5e^{-\frac{4}{3}}$
(C) 3
(D) $2+3e^{-\frac{2}{3}}$
(E) 5

Problem 21 ‡
The cumulative distribution function for health care costs experienced by a policyholder is modeled by the function

$$F(x)=\begin{cases} 1-e^{-\frac{x}{100}}, & \text{for } x>0\\ 0, & \text{otherwise.}\end{cases}$$

The policy has a deductible of 20. An insurer reimburses the policyholder for 100% of health care costs between 20 and 120 less the deductible. Health care costs above 120 are reimbursed at 50%. Let G be the cumulative distribution function of reimbursements given that the reimbursement is positive. Calculate $G(115)$.

(A) 0.683
(B) 0.727
(C) 0.741
(D) 0.757
(E) 0.777

Problem 22 ‡
Let T denote the time in minutes for a customer service representative to respond to 10 telephone inquiries. T is uniformly distributed on the interval with endpoints 8 minutes and 12 minutes. Let R denote the average rate, in customers per minute, at which the representative responds to inquiries.
Find the density function $f_R(r)$ of R.

(A) $\frac{12}{5}$
(B) $3-\frac{5}{2r}$
(C) $3r-\frac{5\ln r}{2}$
(D) $\frac{10}{r^2}$
(E) $\frac{5}{2r^2}$

Problem 23 ‡
An insurance company insures a large number of drivers. Let X be the random variable representing the company's losses under collision insurance, and let Y represent the company's losses under liability insurance. X and Y have joint density function

$$f_{XY}(x,y) = \begin{cases} \frac{2x+2-y}{4} & 0 < x < 1, 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the probability that the total loss is at least 1 ?

(A) 0.33
(B) 0.38
(C) 0.41
(D) 0.71
(E) 0.75

Problem 24 ‡
A device contains two circuits. The second circuit is a backup for the first, so the second is used only when the first has failed. The device fails when and only when the second circuit fails.
Let X and Y be the times at which the first and second circuits fail, respectively. X and Y have joint probability density function

$$f_{XY}(x,y) = \begin{cases} 6e^{-x}e^{-2y} & 0 < x < y < \infty \\ 0 & \text{otherwise.} \end{cases}$$

What is the expected time at which the device fails?

(A) 0.33
(B) 0.50
(C) 0.67
(D) 0.83
(E) 1.50

Problem 25 ‡
A client spends X minutes in an insurance agent's waiting room and Y minutes meeting with the agent. The joint density function of X and Y can be modeled by

$$f(x,y) = \begin{cases} \frac{1}{800}e^{\frac{x}{40}+\frac{y}{20}} & \text{for } x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability that a client spends less than 60 minutes at the agent's office. You do NOT have to evaluate the integrals.

(A) $\frac{1}{800}\int_0^{40}\int_0^{20} e^{-\frac{x}{40}}e^{-\frac{y}{20}}$
(B) $\frac{1}{800}\int_0^{40}\int_0^{20-x} e^{-\frac{x}{40}}e^{-\frac{y}{20}}$
(C) $\frac{1}{800}\int_0^{20}\int_0^{40-x} e^{-\frac{x}{40}}e^{-\frac{y}{20}}$
(D) $\frac{1}{800}\int_0^{60}\int_0^{60} e^{-\frac{x}{40}}e^{-\frac{y}{20}}$
(E) $\frac{1}{800}\int_0^{60}\int_0^{60-x} e^{-\frac{x}{40}}e^{-\frac{y}{20}}$

Problem 26 ‡
An insurance company sells two types of auto insurance policies: Basic and Deluxe. The time until the next Basic Policy claim is an exponential random variable with mean two days. The time until the next Deluxe Policy claim is an independent exponential random variable with mean three days.
What is the probability that the next claim will be a Deluxe Policy claim?

(A) 0.172
(B) 0.223
(C) 0.400
(D) 0.487
(E) 0.500

Problem 27 ‡
A device containing two key components fails when, and only when, both components fail. The lifetimes, X and Y, of these components are independent with common density function $f(t) = e^{-t}, t > 0$. The cost, Z, of operating the device until failure is $2X + Y$.
Find the probability density function of Z.

(A) $e^{-\frac{x}{2}} - e^{-x}$
(B) $2(e^{-\frac{x}{2}} - e^{-x})$
(C) $\frac{x^2e^{-x}}{2}$
(D) $\frac{e^{-\frac{x}{2}}}{2}$
(E) $\frac{e^{-\frac{x}{3}}}{3}$

Problem 28 ‡

Let X and Y be continuous random variables with joint density function

$$f_{XY}(x,y) = \begin{cases} 24xy & 0 < x < 1, 0 < y < 1-x \\ 0 & \text{otherwise.} \end{cases}$$

Calculate $\Pr\left(Y < X | X = \frac{1}{3}\right)$.

(A) $\frac{1}{27}$
(B) $\frac{2}{27}$
(C) $\frac{1}{4}$
(D) $\frac{1}{3}$
(E) $\frac{4}{9}$

Problem 29 ‡
You are given the following information about N, the annual number of claims for a randomly selected insured:

$$\Pr(N=0) = \frac{1}{2}$$
$$\Pr(N=1) = \frac{1}{3}$$
$$\Pr(N>1) = \frac{1}{6}$$

Let S denote the total annual claim amount for an insured. When $N = 1, S$ is exponentially distributed with mean 5 . When $N > 1, S$ is exponentially distributed with mean 8 . Determine $\Pr(4 < S < 8)$.

(A) 0.04
(B) 0.08
(C) 0.12
(D) 0.24
(E) 0.25

Problem 30 ‡
Let T_1 and T_2 represent the lifetimes in hours of two linked components in an electronic device. The joint density function for T_1 and T_2 is uniform over the region defined by $0 \le t_1 \le t_2 \le L$, where L is a positive constant.
Determine the expected value of the sum of the squares of T_1 and T_2.

(A) $\frac{L^2}{3}$
(B) $\frac{L^2}{2}$
(C) $\frac{2L^2}{3}$
(D) $\frac{3L^2}{4}$
(E) L^2

Problem 31 ‡
The profit for a new product is given by $Z = 3X - Y - 5$. X and Y are independent random variables with Var(X) = 1 and Var(Y) = 2.
What is the variance of Z?

(A) 1
(B) 5
(C) 7
(D) 11
(E) 16

Problem 32 ‡
Let X and Y denote the values of two stocks at the end of a five-year period. X is uniformly distributed on the interval (0, 12) . Given $X = x$, Y is uniformly distributed on the interval $(0, x)$.
Determine $Cov(X, Y)$ according to this model.

(A) 0
(B) 4
(C) 6
(D) 12
(E) 24

Problem 33 ‡
The stock prices of two companies at the end of any given year are modeled with random variables X and Y that follow a distribution with joint density function

$$f_{XY}(x,y) = \begin{cases} 2x & 0 < x < 1, x < y < x+1 \\ 0 & \text{otherwise.} \end{cases}$$

What is the conditional variance of Y given that $X = x$?

(A) $\frac{1}{12}$

(B) $\frac{7}{6}$
(C) $\frac{x+1}{2}$
(D) $\frac{x^2-1}{6}$
(E) $\frac{x^2+x+1}{3}$

Problem 34 ‡
A fair die is rolled repeatedly. Let X be the number of rolls needed to obtain a 5 and Y the number of rolls needed to obtain a 6. Calculate $E(X|Y=2)$.

(A) 5.0
(B) 5.2
(C) 6.0
(D) 6.6
(E) 6.8

Problem 35 ‡
The number of hurricanes that will hit a certain house in the next ten years is Poisson distributed with mean 4.
Each hurricane results in a loss that is exponentially distributed with mean 1000. Losses are mutually independent and independent of the number of hurricanes.
Calculate the variance of the total loss due to hurricanes hitting this house in the next ten years.

(A) 4,000,000
(B) 4,004,000
(C) 8,000,000
(D) 16,000,000
(E) 20,000,000

Problem 36 ‡
A company insures homes in three cities, J, K, and L . Since sufficient distance separates the cities, it is reasonable to assume that the losses occurring in these cities are independent.
The moment generating functions for the loss distributions of the cities are:

$$M_J(t) = (1-2t)^{-3}$$
$$M_K(t) = (1-2t)^{-2.5}$$
$$M_L(t) = (1-2t)^{-4.5}$$

Let X represent the combined losses from the three cities. Calculate $E(X^3)$.

(A) 1,320
(B) 2,082
(C) 5,760
(D) 8,000
(E) 10,560

Problem 37 ‡
Let X and Y be identically distributed independent random variables such that the moment generating function of $X+Y$ is

$$M(t) = 0.09e^{-2t} + 0.24e^{-t} + 0.34 + 0.24e^{t} + 0.09e^{2t}, \quad -\infty < t < \infty.$$

Calculate $\Pr(X \leq 0)$.

(A) 0.33
(B) 0.34
(C) 0.50
(D) 0.67
(E) 0.70

Problem 38 ‡
A company manufactures a brand of light bulb with a lifetime in months that is normally distributed with mean 3 and variance 1 . A consumer buys a number of these bulbs with the intention of replacing them successively as they burn out. The light bulbs have independent lifetimes.
What is the smallest number of bulbs to be purchased so that the succession of light bulbs produces light for at least 40 months with probability at least 0.9772?

(A) 14
(B) 16
(C) 20
(D) 40
(E) 55

Answers

1. D
2. E
3. E
4. C
5. B
6. D
7. B
8. A
9. B
10. B
11. E
12. E
13. B
14. C
15. B
16. D
17. E
18. B
19. D
20. D
21. B
22. E
23. D
24. D
25. E
26. C
27. A
28. C
29. C
30. C
31. D
32. C
33. A
34. D
35. C

36. E
37. E
38. B

Sample Exam 3

Problem 1 ‡
A marketing survey indicates that 60% of the population owns an automobile, 30% owns a house, and 20% owns both an automobile and a house. What percentage of the population owns an automobile or a house, but not both?

(A) 0.4
(B) 0.5
(C) 0.6
(D) 0.7
(E) 0.9

Problem 2 ‡
A survey of 100 TV watchers revealed that over the last year:
i) 34 watched CBS.
ii) 15 watched NBC.
iii) 10 watched ABC.
iv) 7 watched CBS and NBC.
v) 6 watched CBS and ABC.
vi) 5 watched NBC and ABC.
vii) 4 watched CBS, NBC, and ABC.
viii) 18 watched HGTV and of these, none watched CBS, NBC, or ABC.
Calculate how many of the 100 TV watchers did not watch any of the four channels (CBS, NBC, ABC or HGTV).

(A) 1

(B) 37
(C) 45
(D) 55
(E) 82

Problem 3 ‡
Among a large group of patients recovering from shoulder injuries, it is found that 22% visit both a physical therapist and a chiropractor, whereas 12% visit neither of these. The probability that a patient visits a chiropractor exceeds by 14% the probability that a patient visits a physical therapist.
Determine the probability that a randomly chosen member of this group visits a physical therapist.

(A) 0.26
(B) 0.38
(C) 0.40
(D) 0.48
(E) 0.62

Problem 4 ‡
The probability that a member of a certain class of homeowners with liability and property coverage will file a liability claim is 0.04, and the probability that a member of this class will file a property claim is 0.10. The probability that a member of this class will file a liability claim but not a property claim is 0.01.
Calculate the probability that a randomly selected member of this class of homeowners will not file a claim of either type.

(A) 0.850
(B) 0.860
(C) 0.864
(D) 0.870
(E) 0.890

Problem 5 ‡
An insurance company examines its pool of auto insurance customers and gathers the following information:

(i) All customers insure at least one car.
(ii) 70% of the customers insure more than one car.
(iii) 20% of the customers insure a sports car.
(iv) Of those customers who insure more than one car, 15% insure a sports car.

Calculate the probability that a randomly selected customer insures exactly one car and that car is not a sports car.

(A) 0.13
(B) 0.21
(C) 0.24
(D) 0.25
(E) 0.30

Problem 6 ‡
Upon arrival at a hospital's emergency room, patients are categorized according to their condition as critical, serious, or stable. In the past year:

(i) 10% of the emergency room patients were critical;
(ii) 30% of the emergency room patients were serious;
(iii) the rest of the emergency room patients were stable;
(iv) 40% of the critical patients died;
(v) 10% of the serious patients died; and
(vi) 1% of the stable patients died.

Given that a patient survived, what is the probability that the patient was categorized as serious upon arrival?

(A) 0.06
(B) 0.29
(C) 0.30
(D) 0.39
(E) 0.64

Problem 7 ‡
The probability that a randomly chosen male has a circulation problem is 0.25 . Males who have a circulation problem are twice as likely to be smokers as those who do not have a circulation problem.
What is the conditional probability that a male has a circulation problem,

given that he is a smoker?

(A) $\frac{1}{4}$
(B) $\frac{1}{3}$
(C) $\frac{2}{5}$
(D) $\frac{1}{2}$
(E) $\frac{2}{3}$

Problem 8 ‡
An actuary studying the insurance preferences of automobile owners makes the following conclusions:
(i) An automobile owner is twice as likely to purchase a collision coverage as opposed to a disability coverage.
(ii) The event that an automobile owner purchases a collision coverage is independent of the event that he or she purchases a disability coverage.
(iii) The probability that an automobile owner purchases both collision and disability coverages is 0.15.
What is the probability that an automobile owner purchases neither collision nor disability coverage?

(A) 0.18
(B) 0.33
(C) 0.48
(D) 0.67
(E) 0.82

Problem 9 ‡
Under an insurance policy, a maximum of five claims may be filed per year by a policyholder. Let p_n be the probability that a policyholder files n claims during a given year, where $n = 0, 1, 2, 3, 4, 5$. An actuary makes the following observations:
(i) $p_n \geq p_{n+1}$ for $0 \leq n \leq 4$
(ii) The difference between p_n and p_{n+1} is the same for $0 \leq n \leq 4$
(iii) Exactly 40% of policyholders file fewer than two claims during a given year.
Calculate the probability that a random policyholder will file more than three claims during a given year.

(A) 0.14
(B) 0.16
(C) 0.27
(D) 0.29
(E) 0.33

Problem 10 ‡
An insurance policy pays 100 per day for up to 3 days of hospitalization and 50 per day for each day of hospitalization thereafter.
The number of days of hospitalization, X, is a discrete random variable with probability function

$$p(k) = \begin{cases} \frac{6-k}{15} & k = 1, 2, 3, 4, 5 \\ 0 & \text{otherwise} \end{cases}$$

Determine the expected payment for hospitalization under this policy.

(A) 123
(B) 210
(C) 220
(D) 270
(E) 367

Problem 11 ‡
A hospital receives 1/5 of its flu vaccine shipments from Company X and the remainder of its shipments from other companies. Each shipment contains a very large number of vaccine vials.
For Company Xs shipments, 10% of the vials are ineffective. For every other company, 2% of the vials are ineffective. The hospital tests 30 randomly selected vials from a shipment and finds that one vial is ineffective.
What is the probability that this shipment came from Company X?

(A) 0.10
(B) 0.14
(C) 0.37
(D) 0.63
(E) 0.86

Problem 12 ‡
Let X represent the number of customers arriving during the morning hours

and let Y represent the number of customers arriving during the afternoon hours at a diner. You are given:
i) X and Y are Poisson distributed.
ii) The first moment of X is less than the first moment of Y by 8.
iii) The second moment of X is 60% of the second moment of Y.
Calculate the variance of Y.

(A) 4
(B) 12
(C) 16
(D) 27
(E) 35

Problem 13 ‡
As part of the underwriting process for insurance, each prospective policyholder is tested for high blood pressure. Let X represent the number of tests completed when the first person with high blood pressure is found. The expected value of X is 12.5.
Calculate the probability that the sixth person tested is the first one with high blood pressure.

(A) 0.000
(B) 0.053
(C) 0.080
(D) 0.316
(E) 0.394

Problem 14 ‡
A group insurance policy covers the medical claims of the employees of a small company. The value, V, of the claims made in one year is described by

$$V = 100000Y$$

where Y is a random variable with density function

$$f(x) = \begin{cases} k(1-y)^4 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.
What is the conditional probability that V exceeds 40,000, given that V exceeds 10,000?

(A) 0.08
(B) 0.13
(C) 0.17
(D) 0.20
(E) 0.51

Problem 15 ‡
An insurance policy reimburses a loss up to a benefit limit of 10 . The policyholder's loss, X, follows a distribution with density function:

$$f(x) = \begin{cases} \frac{2}{x^3} & x > 1 \\ 0 & \text{otherwise.} \end{cases}$$

What is the expected value of the benefit paid under the insurance policy?

(A) 1.0
(B) 1.3
(C) 1.8
(D) 1.9
(E) 2.0

Problem 16 ‡
An auto insurance company insures an automobile worth 15,000 for one year under a policy with a 1,000 deductible. During the policy year there is a 0.04 chance of partial damage to the car and a 0.02 chance of a total loss of the car. If there is partial damage to the car, the amount X of damage (in thousands) follows a distribution with density function

$$f(x) = \begin{cases} 0.5003e^{-0.5x} & 0 < x < 15 \\ 0 & \text{otherwise.} \end{cases}$$

What is the expected claim payment?

(A) 320
(B) 328
(C) 352
(D) 380
(E) 540

Problem 17 ‡
An insurance policy on an electrical device pays a benefit of 4000 if the device fails during the first year. The amount of the benefit decreases by 1000 each successive year until it reaches 0 . If the device has not failed by the beginning of any given year, the probability of failure during that year is 0.4. What is the expected benefit under this policy?

(A) 2234
(B) 2400
(C) 2500
(D) 2667
(E) 2694

Problem 18 ‡
An insurance policy is written to cover a loss, X, where X has a uniform distribution on [0, 1000].
At what level must a deductible be set in order for the expected payment to be 25% of what it would be with no deductible?

(A) 250
(B) 375
(C) 500
(D) 625
(E) 750

Problem 19 ‡
Ten years ago at a certain insurance company, the size of claims under homeowner insurance policies had an exponential distribution. Furthermore, 25% of claims were less than \$1000. Today, the size of claims still has an exponential distribution but, owing to inflation, every claim made today is twice the size of a similar claim made 10 years ago. Determine the probability that a claim made today is less than \$1000.

(A) 0.063
(B) 0.125
(C) 0.134
(D) 0.163
(E) 0.250

Problem 20 ‡
A piece of equipment is being insured against early failure. The time from purchase until failure of the equipment is exponentially distributed with mean 10 years. The insurance will pay an amount x if the equipment fails during the first year, and it will pay $0.5x$ if failure occurs during the second or third year. If failure occurs after the first three years, no payment will be made.
At what level must x be set if the expected payment made under this insurance is to be 1000 ?

(A) 3858
(B) 4449
(C) 5382
(D) 5644
(E) 7235

Problem 21 ‡
The time, T, that a manufacturing system is out of operation has cumulative distribution function

$$F(t) = \begin{cases} 1 - \left(\frac{2}{t}\right)^2 & t > 2 \\ 0 & \text{otherwise} \end{cases}$$

The resulting cost to the company is $Y = T^2$. Determine the density function of Y, for $y > 4$.

(A) $\frac{4}{y^2}$
(B) $\frac{8}{y^{\frac{3}{2}}}$
(C) $\frac{8}{y^3}$
(D) $\frac{16}{y}$
(E) $\frac{1024}{y^5}$

Problem 22 ‡
The monthly profit of Company A can be modeled by a continuous random variable with density function f_A. Company B has a monthly profit that is twice that of Company A.
Determine the probability density function of the monthly profit of Company B.

(A) $\frac{1}{2}f\left(\frac{x}{2}\right)$
(B) $f\left(\frac{x}{2}\right)$
(C) $2f\left(\frac{x}{2}\right)$
(D) $2f(x)$
(E) $2f(2x)$

Problem 23 ‡
A car dealership sells 0, 1, or 2 luxury cars on any day. When selling a car, the dealer also tries to persuade the customer to buy an extended warranty for the car. Let X denote the number of luxury cars sold in a given day, and let Y denote the number of extended warranties sold. Given the following information

$$\begin{aligned}
\Pr(X=0,Y=0)=&\frac{1}{6}\\
\Pr(X=1,Y=0)=&\frac{1}{12}\\
\Pr(X=1,Y=1)=&\frac{1}{6}\\
\Pr(X=2,Y=0)=&\frac{1}{12}\\
\Pr(X=2,Y=1)=&\frac{1}{3}\\
\Pr(X=2,Y=2)=&\frac{1}{6}
\end{aligned}$$

What is the variance of X?

(A) 0.47
(B) 0.58
(C) 0.83
(D) 1.42
(E) 2.58

Problem 24 ‡
The future lifetimes (in months) of two components of a machine have the following joint density function:

$$f_{XY}(x,y)=\begin{cases}\frac{6}{125000}(50-x-y) & 0<x<50-y<50\\ 0 & \text{otherwise.}\end{cases}$$

What is the probability that both components are still functioning 20 months from now?

(A) $\frac{6}{125,000}\int_0^{20}\int_0^{20}(50-x-y)dydx$

(B) $\frac{6}{125,000}\int_{20}^{30}\int_{02}^{50-x}(50-x-y)dydx$

(C) $\frac{6}{125,000}\int_{20}^{30}\int_{20}^{50-x-y}(50-x-y)dydx$

(D) $\frac{6}{125,000}\int_{20}^{50}\int_{20}^{50-x}(50-x-y)dydx$

(E) $\frac{6}{125,000}\int_{20}^{50}\int_{20}^{50-x-y}(50-x-y)dydx$

Problem 25 ‡
Automobile policies are separated into two groups: low-risk and high-risk. Actuary Rahul examines low-risk policies, continuing until a policy with a claim is found and then stopping. Actuary Toby follows the same procedure with high-risk policies. Each low-risk policy has a 10% probability of having a claim. Each high-risk policy has a 20% probability of having a claim. The claim statuses of polices are mutually independent.
Calculate the probability that Actuary Rahul examines fewer policies than Actuary Toby.

(A) 0.2857
(B) 0.3214
(C) 0.3333
(D) 0.3571
(E) 0.4000

Problem 26 ‡
Two insurers provide bids on an insurance policy to a large company. The bids must be between 2000 and 2200 . The company decides to accept the lower bid if the two bids differ by 20 or more. Otherwise, the company will consider the two bids further. Assume that the two bids are independent and are both uniformly distributed on the interval from 2000 to 2200.
Determine the probability that the company considers the two bids further.

(A) 0.10

(B) 0.19
(C) 0.20
(D) 0.41
(E) 0.60

Problem 27 ‡
A company offers earthquake insurance. Annual premiums are modeled by an exponential random variable with mean 2. Annual claims are modeled by an exponential random variable with mean 1. Premiums and claims are independent. Let X denote the ratio of claims to premiums.
What is the density function of X?

(A) $\frac{1}{2x+1}$
(B) $\frac{2}{(2x+1)^2}$
(C) e^{-x}
(D) $2e^{-2x}$
(E) xe^{-x}

Problem 28 ‡
Once a fire is reported to a fire insurance company, the company makes an initial estimate, X, of the amount it will pay to the claimant for the fire loss. When the claim is finally settled, the company pays an amount, Y, to the claimant. The company has determined that X and Y have the joint density function

$$f_{XY}(x,y) = \begin{cases} \frac{2}{x^2(x-1)}y^{-(2x-1)/(x-1)} & x > 1, y > 1 \\ 0 & \text{otherwise.} \end{cases}$$

Given that the initial claim estimated by the company is 2, determine the probability that the final settlement amount is between 1 and 3 .

(A) $\frac{1}{9}$
(B) $\frac{2}{9}$
(C) $\frac{1}{3}$
(D) $\frac{2}{3}$
(E) $\frac{8}{9}$

Problem 29 ‡
The distribution of Y, given X, is uniform on the interval $[0, X]$. The marginal

density of X is

$$f_X(x) = \begin{cases} 2x & \text{for } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Determine the conditional density of X, given $Y = y > 0$.

(A) 1
(B) 2
(C) $2x$
(D) $\frac{1}{y}$
(E) $\frac{1}{1-y}$

Problem 30 ‡
A machine consists of two components, whose lifetimes have the joint density function

$$f(x,y) = \begin{cases} \frac{1}{50} & \text{for } x > 0, y > 0, x + y < 10 \\ 0 & \text{otherwise.} \end{cases}$$

The machine operates until both components fail. Calculate the expected operational time of the machine.

(A) 1.7
(B) 2.5
(C) 3.3
(D) 5.0
(E) 6.7

Problem 31 ‡
A company has two electric generators. The time until failure for each generator follows an exponential distribution with mean 10. The company will begin using the second generator immediately after the first one fails.
What is the variance of the total time that the generators produce electricity?

(A) 10
(B) 20
(C) 50
(D) 100
(E) 200

Problem 32 ‡

Let X denote the size of a surgical claim and let Y denote the size of the associated hospital claim. An actuary is using a model in which $E(X) = 5, E(X^2) = 27.4, E(Y) = 7, E(Y^2) = 51.4$, and $Var(X+Y) = 8$.

Let $C_1 = X+Y$ denote the size of the combined claims before the application of a 20% surcharge on the hospital portion of the claim, and let C_2 denote the size of the combined claims after the application of that surcharge.

Calculate $Cov(C_1, C_2)$.

(A) 8.80
(B) 9.60
(C) 9.76
(D) 11.52
(E) 12.32

Problem 33 ‡

An actuary determines that the annual numbers of tornadoes in counties P and Q are jointly distributed as follows:

$X\backslash Y$	0	1	2	$P_X(x)$
0	0.12	0.13	0.05	0.30
1	0.06	0.15	0.15	0.36
2	0.05	0.12	0.10	0.27
3	0.02	0.03	0.02	0.07
$p_Y(y)$	0.25	0.43	0.32	1

where X is the number of tornadoes in county Q and Y that of county P.

Calculate the conditional variance of the annual number of tornadoes in county Q, given that there are no tornadoes in county P.

(A) 0.51
(B) 0.84
(C) 0.88
(D) 0.99
(E) 1.76

Problem 34 ‡

A driver and a passenger are in a car accident. Each of them independently has probability 0.3 of being hospitalized. When a hospitalization occurs, the

loss is uniformly distributed on [0, 1]. When two hospitalizations occur, the losses are independent.
Calculate the expected number of people in the car who are hospitalized, given that the total loss due to hospitalizations from the accident is less than 1.

(A) 0.510
(B) 0.534
(C) 0.600
(D) 0.628
(E) 0.800

Problem 35 ‡
An insurance company insures two types of cars, economy cars and luxury cars. The damage claim resulting from an accident involving an economy car has normal $N(7, 1)$ distribution, the claim from a luxury car accident has normal $N(20, 6)$ distribution.
Suppose the company receives three claims from economy car accidents and one claim from a luxury car accident. Assuming that these four claims are mutually independent, what is the probability that the total claim amount from the three economy car accidents exceeds the claim amount from the luxury car accident?

(A) 0.731
(B) 0.803
(C) 0.629
(D) 0.235
(E) 0.296

Problem 36 ‡
Let X_1, X_2, X_3 be independent discrete random variables with common probability mass function

$$\Pr(x) = \begin{cases} \frac{1}{3} & x = 0 \\ \frac{2}{3} & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Determine the moment generating function $M(t)$, of $Y = X_1 X_2 X_3$.

(A) $\frac{19}{27} + \frac{8}{27}e^t$
(B) $1 + 2e^t$
(C) $\left(\frac{1}{3} + \frac{2}{3}e^t\right)^3$
(D) $\frac{1}{27} + \frac{8}{27}e^{3t}$
(E) $\frac{1}{3} + \frac{2}{3}e^{3t}$

Problem 37 ‡
In an analysis of healthcare data, ages have been rounded to the nearest multiple of 5 years. The difference between the true age and the rounded age is assumed to be uniformly distributed on the interval from -2.5 years to 2.5 years. The healthcare data are based on a random sample of 48 people.
What is the approximate probability that the mean of the rounded ages is within 0.25 years of the mean of the true ages?

(A) 0.14
(B) 0.38
(C) 0.57
(D) 0.77
(E) 0.88

Problem 38 ‡
Let X and Y be the number of hours that a randomly selected person watches movies and sporting events, respectively, during a three-month period. The following information is known about X and Y :

$$\begin{array}{ll} \mathrm{E(X)} = & 50 \\ \mathrm{E(Y)} = & 20 \\ \mathrm{Var(X)} = & 50 \\ \mathrm{Var(Y)} = & 30 \\ \mathrm{Cov\ (X,Y)} = & 10 \end{array}$$

One hundred people are randomly selected and observed for these three months. Let T be the total number of hours that these one hundred people watch movies or sporting events during this three-month period.
Approximate the value of $\Pr(T < 7100)$.

(A) 0.62
(B) 0.84
(C) 0.87

(D) 0.92
(E) 0.97

Answers

1. B
2. B
3. D
4. E
5. B
6. B
7. C
8. B
9. C
10. C
11. A
12. E
13. B
14. B
15. D
16. B
17. E
18. C
19. C
20. D
21. A
22. A
23. B
24. D
25. A
26. B
27. B
28. E
29. E
30. D
31. E
32. A
33. D
34. B
35. C

36. A
37. D
38. B

Sample Exam 4

Problem 1 ‡
35% of visits to a primary care physicians (PCP) office results in neither lab work nor referral to a specialist. Of those coming to a PCPs office, 30% are referred to specialists and 40% require lab work.
What percentage of visit to a PCPs office results in both lab work and referral to a specialist?

(A) 0.05
(B) 0.12
(C) 0.18
(D) 0.25
(E) 0.35

Problem 2 ‡
Thirty items are arranged in a 6-by-5 array as shown.

A_1	A_2	A_3	A_4	A_5
A_6	A_7	A_8	A_9	A_{10}
A_{11}	A_{12}	A_{13}	A_{14}	A_{15}
A_{16}	A_{17}	A_{18}	A_{19}	A_{20}
A_{21}	A_{22}	A_{23}	A_{24}	A_{25}
A_{26}	A_{27}	A_{28}	A_{29}	A_{30}

Calculate the number of ways to form a set of three distinct items such that no two of the selected items are in the same row or same column.

(A) 200

(B) 760
(C) 1200
(D) 4560
(E) 7200

Problem 3 ‡
In modeling the number of claims filed by an individual under an automobile policy during a three-year period, an actuary makes the simplifying assumption that for all integers $n \geq 0, p_{n+1} = \frac{1}{5}p_n$, where p_n represents the probability that the policyholder files n claims during the period.
Under this assumption, what is the probability that a policyholder files more than one claim during the period?

(A) 0.04
(B) 0.16
(C) 0.20
(D) 0.80
(E) 0.96

Problem 4 ‡
A store has 80 modems in its inventory, 30 coming from Source A and the remainder from Source B. Of the modems from Source A, 20% are defective. Of the modems from Source B, 8% are defective.
Calculate the probability that exactly two out of a random sample of five modems from the store's inventory are defective.

(A) 0.010
(B) 0.078
(C) 0.102
(D) 0.105
(E) 0.125

Problem 5 ‡
An actuary is studying the prevalence of three health risk factors, denoted by A, B, and C, within a population of women. For each of the three factors, the probability is 0.1 that a woman in the population has only this risk factor (and no others). For any two of the three factors, the probability is 0.12 that she has exactly these two risk factors (but not the other). The probability

that a woman has all three risk factors, given that she has A and B, is $\frac{1}{3}$. What is the probability that a woman has none of the three risk factors, given that she does not have risk factor A?

(A) 0.280
(B) 0.311
(C) 0.467
(D) 0.484
(E) 0.700

Problem 6 ‡
A health study tracked a group of persons for five years. At the beginning of the study, 20% were classified as heavy smokers, 30% as light smokers, and 50% as nonsmokers.
Results of the study showed that light smokers were twice as likely as nonsmokers to die during the five-year study, but only half as likely as heavy smokers.
A randomly selected participant from the study died over the five-year period. Calculate the probability that the participant was a heavy smoker.

(A) 0.20
(B) 0.25
(C) 0.35
(D) 0.42
(E) 0.57

Problem 7 ‡
A study of automobile accidents produced the following data:

Model year	Proportion of all vehicles	Probability of involvement in an accident
1997	0.16	0.05
1998	0.18	0.02
1999	0.20	0.03
Other	0.46	0.04

An automobile from one of the model years 1997, 1998, and 1999 was involved in an accident. Determine the probability that the model year of this automobile is 1997.

(A) 0.22
(B) 0.30
(C) 0.33
(D) 0.45
(E) 0.50

Problem 8 ‡
An insurance company pays hospital claims. The number of claims that include emergency room or operating room charges is 85% of the total number of claims. The number of claims that do not include emergency room charges is 25% of the total number of claims. The occurrence of emergency room charges is independent of the occurrence of operating room charges on hospital claims.
Calculate the probability that a claim submitted to the insurance company includes operating room charges.

(A) 0.10
(B) 0.20
(C) 0.25
(D) 0.40
(E) 0.80

Problem 9 ‡
Suppose that an insurance company has broken down yearly automobile claims for drivers from age 16 through 21 as shown in the following table.

Amount of claim	Probability
$ 0	0.80
$ 2000	0.10
$ 4000	0.05
$ 6000	0.03
$ 8000	0.01
$ 10000	0.01

How much should the company charge as its average premium in order to break even on costs for claims?

(A) 706
(B) 760
(C) 746
(D) 766
(E) 700

Problem 10 ‡
An insurance company sells a one-year automobile policy with a deductible of 2 . The probability that the insured will incur a loss is 0.05 . If there is a loss, the probability of a loss of amount N is $\frac{K}{N}$, for $N = 1, \cdots, 5$ and K a constant. These are the only possible loss amounts and no more than one loss can occur.
Determine the net premium for this policy.

(A) 0.031
(B) 0.066
(C) 0.072
(D) 0.110
(E) 0.150

Problem 11 ‡
A company establishes a fund of 120 from which it wants to pay an amount, C, to any of its 20 employees who achieve a high performance level during the coming year. Each employee has a 2% chance of achieving a high performance level during the coming year, independent of any other employee.
Determine the maximum value of C for which the probability is less than 1% that the fund will be inadequate to cover all payments for high performance.

(A) 24
(B) 30
(C) 40
(D) 60
(E) 120

Problem 12 ‡
An actuary has discovered that policyholders are three times as likely to file two claims as to file four claims.
If the number of claims filed has a Poisson distribution, what is the variance

of the number of claims filed?

(A) $\frac{1}{\sqrt{3}}$
(B) 1
(C) $\sqrt{2}$
(D) 2
(E) 4

Problem 13 ‡
A company takes out an insurance policy to cover accidents that occur at its manufacturing plant. The probability that one or more accidents will occur during any given month is $\frac{3}{5}$.
The number of accidents that occur in any given month is independent of the number of accidents that occur in all other months.
Calculate the probability that there will be at least four months in which no accidents occur before the fourth month in which at least one accident occurs.

(A) 0.01
(B) 0.12
(C) 0.23
(D) 0.29
(E) 0.41

Problem 14 ‡
The loss due to a fire in a commercial building is modeled by a random variable X with density function

$$f(x) = \begin{cases} 0.005(20-x) & 0 < x < 20 \\ 0 & \text{otherwise} \end{cases}$$

Given that a fire loss exceeds 8, what is the probability that it exceeds 16 ?

(A) $\frac{1}{25}$
(B) $\frac{1}{9}$
(C) $\frac{1}{8}$
(D) $\frac{1}{3}$
(E) $\frac{3}{7}$

Problem 15 ‡
Claim amounts for wind damage to insured homes are independent random variables with common density function

$$f(x) = \begin{cases} \frac{3}{x^4} & x > 1 \\ 0 & \text{otherwise} \end{cases}$$

where x is the amount of a claim in thousands.
Suppose 3 such claims will be made, what is the expected value of the largest of the three claims?

(A) 2025
(B) 2700
(C) 3232
(D) 3375
(E) 4500

Problem 16 ‡
An insurance company's monthly claims are modeled by a continuous, positive random variable X, whose probability density function is proportional to $(1+x)^{-4}$, where $0 < x < \infty$ and 0 otherwise.
Determine the company's expected monthly claims.

(A) $\frac{1}{6}$
(B) $\frac{1}{3}$
(C) $\frac{1}{2}$
(D) 1
(E) 3

Problem 17 ‡
A man purchases a life insurance policy on his 40^{th} birthday. The policy will pay 5000 only if he dies before his 50^{th} birthday and will pay 0 otherwise. The length of lifetime, in years, of a male born the same year as the insured has the cumulative distribution function

$$F(t) = \begin{cases} 1 - e^{\frac{1-1.1^t}{1000}}, & t > 0 \\ 0 & t \leq 0 \end{cases}$$

Calculate the expected payment to the man under this policy.

(A) 333
(B) 348
(C) 421
(D) 549
(E) 574

Problem 18 ‡
The warranty on a machine specifies that it will be replaced at failure or age 4, whichever occurs first. The machine's age at failure, X, has density function

$$f(x) = \begin{cases} \frac{1}{5} & 0 < x < 5 \\ 0 & \text{otherwise} \end{cases}$$

Let Y be the age of the machine at the time of replacement. Determine the variance of Y.

(A) 1.3
(B) 1.4
(C) 1.7
(D) 2.1
(E) 7.5

Problem 19 ‡
The number of days that elapse between the beginning of a calendar year and the moment a high-risk driver is involved in an accident is exponentially distributed. An insurance company expects that 30% of high-risk drivers will be involved in an accident during the first 50 days of a calendar year.
What portion of high-risk drivers are expected to be involved in an accident during the first 80 days of a calendar year?

(A) 0.15
(B) 0.34
(C) 0.43
(D) 0.57
(E) 0.66

Problem 20 ‡
An insurance policy reimburses dental expense, X, up to a maximum benefit

of 250 . The probability density function for X is:

$$f(x) = \begin{cases} ce^{-0.004x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where c is a constant. Calculate the median benefit for this policy.

(A) 161
(B) 165
(C) 173
(D) 182
(E) 250

Problem 21 ‡
An investment account earns an annual interest rate R that follows a uniform distribution on the interval $(0.04, 0.08)$. The value of a 10,000 initial investment in this account after one year is given by $V = 10,000e^R$.
Determine the cumulative distribution function, $F_V(v)$ of V.

(A) $\frac{10{,}000e^{\frac{v}{10{,}000}}-10{,}408}{425}$
(B) $25e^{\frac{v}{10{,}000}} - 0.04$
(C) $\frac{v-10{,}408}{10{,}833-10{,}408}$
(D) $\frac{25}{v}$
(E) $25\left[\ln\left(\frac{v}{10{,}000}\right) - 0.04\right]$

Problem 22 ‡
A company is reviewing tornado damage claims under a farm insurance policy. Let X be the portion of a claim representing damage to the house and let Y be the portion of the same claim representing damage to the rest of the property. The joint density function of X and Y is

$$f_{XY}(x,y) = \begin{cases} 6[1-(x+y)] & x > 0, y > 0, x+y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Determine the probability that the portion of a claim representing damage to the house is less than 0.2.

(A) 0.360

(B) 0.480
(C) 0.488
(D) 0.512
(E) 0.520

Problem 23 ‡
Let X and Y be continuous random variables with joint density function

$$f_{XY}(x,y)=\begin{cases} 15y & x^2\le y\le x\\ 0 & \text{otherwise.}\end{cases}$$

Find the marginal density function of Y.

(A)

$$g(y)=\begin{cases} 15y & 0<y<1\\ 0 & \text{otherwise}\end{cases}$$

(B)

$$g(y)=\begin{cases} \frac{15y^2}{2} & x^2<y<x\\ 0 & \text{otherwise}\end{cases}$$

(C)

$$g(y)=\begin{cases} \frac{15y^2}{2} & 0<y<1\\ 0 & \text{otherwise}\end{cases}$$

(D)

$$g(y)=\begin{cases} 15y^{\frac{3}{2}}(1-y^{\frac{1}{2}}) & x^2<y<x\\ 0 & \text{otherwise}\end{cases}$$

(E)

$$g(y)=\begin{cases} 15y^{\frac{3}{2}}(1-y^{\frac{1}{2}}) & 0<y<1\\ 0 & \text{otherwise}\end{cases}$$

Problem 24 ‡
Let X and Y be random losses with joint density function

$$f_{XY}(x,y) = e^{-(x+y)}, \quad x > 0, y > 0$$

and 0 otherwise. An insurance policy is written to reimburse $X + Y$.
Calculate the probability that the reimbursement is less than 1.

(A) e^{-2}
(B) e^{-1}
(C) $1 - e^{-1}$
(D) $1 - 2e^{-1}$
(E) $1 - 2e^{-2}$

Problem 25 ‡
A study is being conducted in which the health of two independent groups of ten policyholders is being monitored over a one-year period of time. Individual participants in the study drop out before the end of the study with probability 0.2 (independently of the other participants).
What is the probability that at least 9 participants complete the study in one of the two groups, but not in both groups?

(A) 0.096
(B) 0.192
(C) 0.235
(D) 0.376
(E) 0.469

Problem 26 ‡
A family buys two policies from the same insurance company. Losses under the two policies are independent and have continuous uniform distributions on the interval from 0 to 10. One policy has a deductible of 1 and the other has a deductible of 2. The family experiences exactly one loss under each policy.
Calculate the probability that the total benefit paid to the family does not exceed 5.

(A) 0.13
(B) 0.25

(C) 0.30
(D) 0.32
(E) 0.42

Problem 27 ‡
An insurance company determines that N, the number of claims received in a week, is a random variable with $P[N = n] = \frac{1}{2^{n+1}}$, where $n \geq 0$. The company also determines that the number of claims received in a given week is independent of the number of claims received in any other week.
Determine the probability that exactly seven claims will be received during a given two-week period.

(A) $\frac{1}{256}$
(B) $\frac{1}{128}$
(C) $\frac{7}{512}$
(D) $\frac{1}{64}$
(E) $\frac{1}{32}$

Problem 28 ‡
A company offers a basic life insurance policy to its employees, as well as a supplemental life insurance policy. To purchase the supplemental policy, an employee must first purchase the basic policy.
Let X denote the proportion of employees who purchase the basic policy, and Y the proportion of employees who purchase the supplemental policy. Let X and Y have the joint density function $f_{XY}(x, y) = 2(x + y)$ on the region where the density is positive.
Given that 10% of the employees buy the basic policy, what is the probability that fewer than 5% buy the supplemental policy?

(A) 0.010
(B) 0.013
(C) 0.108
(D) 0.417
(E) 0.500

Problem 29 ‡
An insurance policy is written to cover a loss X where X has density function

$$f_X(x) = \begin{cases} \frac{3}{8}x^2 & 0 \leq x \leq 2 \\ 0 & \text{otherwise.} \end{cases}$$

The time T (in hours) to process a claim of size x, where $0 \leq x \leq 2$, is uniformly distributed on the interval from x to $2x$.
Calculate the probability that a randomly chosen claim on this policy is processed in three hours or more.

(A) 0.17
(B) 0.25
(C) 0.32
(D) 0.58
(E) 0.83

Problem 30 ‡
The profit for a new product is given by $Z = 3X - Y - 5$, where X and Y are independent random variables with $\text{Var}(X) = 1$ and $\text{Var}(Y) = 2$. What is the variance of Z?

(A) 1
(B) 5
(C) 7
(D) 11
(E) 16

Problem 31 ‡
A joint density function is given by

$$f_{XY}(x,y) = \begin{cases} kx & 0 < x, y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $\text{Cov}(X,Y)$

(A) $-\frac{1}{6}$
(B) 0
(C) $\frac{1}{9}$
(D) $\frac{1}{6}$
(E) $\frac{2}{3}$

Problem 32 ‡
Claims filed under auto insurance policies follow a normal distribution with mean 19,400 and standard deviation 5,000.

What is the probability that the average of 25 randomly selected claims exceeds 20,000 ?

(A) 0.01
(B) 0.15
(C) 0.27
(D) 0.33
(E) 0.45

Problem 33 ‡
The joint probability density for X and Y is

$$f(x,y) = \begin{cases} 2e^{-(x+2y)}, & \text{for } x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the variance of Y given that $X > 3$ and $Y > 3$.

(A) 0.25
(B) 0.50
(C) 1.00
(D) 3.25
(E) 3.50

Problem 34 ‡
New dental and medical plan options will be offered to state employees next year. An actuary uses the following density function to model the joint distribution of the proportion X of state employees who will choose Dental Option 1 and the proportion Y who will choose Medical Option 1 under the new plan options:

$$f(x,y) = \begin{cases} 0.50 & \text{for } 0 < x, y < 0.5 \\ 1.25 & \text{for } 0 < x < 0.5, 0.5 < y < 1 \\ 1.50 & \text{for } 0.5 < x < 1, 0 < y < 0.5 \\ 0.75 & \text{for } 0.5 < x < 1, 0.5 < y < 1. \end{cases}$$

Calculate $\text{Var}(Y|X = 0.75)$.

(A) 0.000
(B) 0.061

(C) 0.076
(D) 0.083
(E) 0.141

Problem 35 ‡
X and Y are independent random variables with common moment generating function $M(t) = e^{\frac{t^2}{2}}$. Let $W = X + Y$ and $Z = X - Y$. Determine the joint moment generating function, $M(t_1, t_2)$ of W and Z.

(A) $e^{2t_1^2+2t_2^2}$
(B) $e^{(t_1-t_2)^2}$
(C) $e^{(t_1+t_2)^2}$
(D) $e^{2t_1t_2}$
(E) $e^{t_1^2+t_2^2}$

Problem 36 ‡
Two instruments are used to measure the height, h, of a tower. The error made by the less accurate instrument is normally distributed with mean 0 and standard deviation $0.0056h$. The error made by the more accurate instrument is normally distributed with mean 0 and standard deviation $0.0044h$. Assuming the two measurements are independent random variables, what is the probability that their average value is within $0.005h$ of the height of the tower?

(A) 0.38
(B) 0.47
(C) 0.68
(D) 0.84
(E) 0.90

Problem 37 ‡
A charity receives 2025 contributions. Contributions are assumed to be independent and identically distributed with mean 3125 and standard deviation 250.
Calculate the approximate 90th percentile for the distribution of the total contributions received.

(A) 6,328,000

(B) 6,338,000
(C) 6,343,000
(D) 6,784,000
(E) 6,977,000

Problem 38 ‡
The total claim amount for a health insurance policy follows a distribution with density function

$$f(x) = \begin{cases} \frac{1}{1000}e^{-\frac{x}{1000}} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

The premium for the policy is set at 100 over the expected total claim amount. If 100 policies are sold, what is the approximate probability that the insurance company will have claims exceeding the premiums collected?

(A) 0.001
(B) 0.159
(C) 0.333
(D) 0.407
(E) 0.460

Answers

1. A
2. C
3. A
4. C
5. C
6. D
7. D
8. D
9. B
10. A
11. D
12. D
13. D
14. B
15. A
16. C
17. B
18. C
19. C
20. C
21. E
22. C
23. D
24. D
25. E
26. C
27. D
28. D
29. A
30. D
31. B
32. C
33. A
34. C
35. E

36. D
37. C
38. B

Answer Keys

Section 1

1.1 $A = \{2, 3, 5\}$
1.2 (a)$S = \{TTT, TTH, THT, THH, HTT, HTH, HHT, HHH\}$
(b)$E = \{TTT, TTH, HTT, THT\}$
(c)$F = \{x : x$ is an element of S with more than one head$\}$
1.3 $F \subset E$
1.4 $E = \emptyset$
1.5 (a) Since every element of A is in A, $A \subseteq A$.
(b) Since every element in A is in B and every element in B is in A, $A = B$.
(c) If x is in A then x is in B since $A \subseteq B$. But $B \subseteq C$ and this implies that x is in C. Hence, every element of A is also in C. This shows that $A \subseteq C$
1.6 The result is true for $n = 1$ since $1 = \frac{1(1+1)}{2}$. Assume that the equality is true for $1, 2, \cdots, n$. Then

$$\begin{aligned}1 + 2 + \cdots + n + 1 =&(1 + 2 + \cdots + n) + n + 1\\ =&\frac{n(n+1)}{2} + n + 1 = (n+1)[\frac{n}{2} + 1]\\ =&\frac{(n+1)(n+2)}{2}\end{aligned}$$

1.7 Let $S_n = 1^2 + 2^2 + 3^2 + \cdots + n^2$. For $n = 1$, we have $S_1 = 1 = \frac{1(1+1)(2+1)}{6}$. Suppose that $S_n = \frac{n(n+1)(2n+1)}{6}$. We next want to show that $S_{n+1} = \frac{(n+1)(n+2)(2n+3)}{6}$. Indeed, $S_{n+1} = 1^2 + 2^2 + 3^2 + \cdots + n^2 + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + (n+1)^2 = (n+1)\left[\frac{n(2n+1)}{6} + n + 1\right] = \frac{(n+1)(n+2)(2n+3)}{6}$

1.8 The result is true for $n = 1$. Suppose true up to n. Then

$$\begin{aligned}(1+x)^{n+1} =&(1+x)(1+x)^n\\ \geq&(1+x)(1+nx), \quad \textit{since } 1+x>0\\ =&1+nx+x+nx^2\\ =&1+nx^2+(n+1)x \geq 1+(n+1)x\end{aligned}$$

1.9 The identity is valid for $n = 1$. Assume true for $1, 2, \cdots, n$. Then

$$\begin{aligned}1+a+a^2+\cdots+a^n =&[1+a+a^2+\cdots+a^{n-1}]+a^n\\ =&\frac{1-a^n}{1-a}+a^n = \frac{1-a^{n+1}}{1-a} \blacksquare\end{aligned}$$

1.10 (a) 55 sandwiches with tomatoes or onions.
(b) There are 40 sandwiches with onions.
(c) There are 10 sandwiches with onions but not tomatoes

1.11 (a) 20 (b) 5 (c) 11 (d) 42 (e) 46 (f) 46
1.12 Since We have

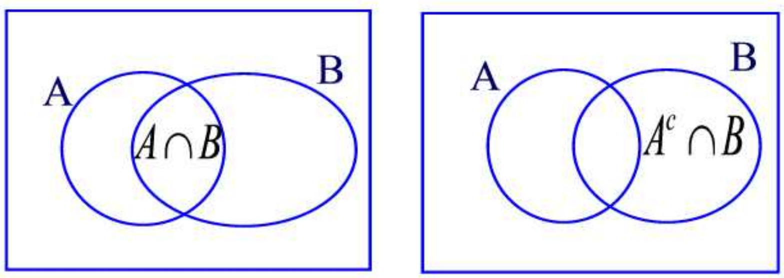
$$S = \{(H,1),(H,2),(H,3),(H,4),(H,5),(H,6),(T,H),(T,T)\}$$

and $n(S) = 8$.
1.13 Suppose that $f(a) = f(b)$. Then $3a+5 = 3b+5 \Longrightarrow 3a+5-5 = 3b+5-5 \Longrightarrow 3a = 3b \Longrightarrow \frac{3a}{3} = \frac{3b}{3} \Longrightarrow a = b$. That is, f is one-to-one.
Let $y \in \mathbb{R}$. From the equation $y = 3x+5$ we find $x = \frac{y-5}{3} \in \mathbb{R}$ and $f(x) = f\left(\frac{3y-5}{3}\right) = y$. That is, f is onto.
1.14 5
1.15 (a) The condition $f(n) = f(m)$ with n even and m odd leads to $n+m = 1$ with $n, m \in \mathbb{N}$ which cannot happen.
(b) Suppose that $f(n) = f(m)$. If n and m are even, we have $\frac{n}{2} = \frac{m}{2} \Longrightarrow n = m$. If n and m are odd then $-\frac{n-1}{2} = -\frac{m-1}{2} \Longrightarrow n = m$. Thus, f is one-to-one. Now, if $m = 0$ then $n = 1$ and $f(n) = m$. If $m \in \mathbb{N} = \mathbb{Z}^+$ then $n = 2m$ and $f(n) = m$. If $n \in \mathbb{Z}^-$ then $n = 2|m|+1$ and $f(n) = m$. Thus, f is onto. If follows that $\mathbb{Z}$ is countable.
1.16 Suppose the contrary. That is, there is a $b \in A$ such that $f(b) = B$. Since $B \subseteq A$, either $b \in B$ or $b \notin B$. If $b \in B$ then $b \notin f(b)$. But $B = f(b)$ so

$b \in B$ implies $b \in f(b)$, a contradiction. If $b \notin B$ then $b \in f(b) = B$ which is again a contradiction. Hence, we conclude that there is no onto map from A to its power set.
1.17 By the previous problem there is no onto map from $\mathbb{N}$ to $\mathcal{P}(\mathbb{N})$ so that $\mathcal{P}(\mathbb{N})$ is uncountable.

Section 2

2.1



2.2 Since $A \subseteq B$, we have $A \cup B = B$. Now the result follows from the previous problem.
2.3 Let

G = event that a viewer watched gymnastics
B = event that a viewer watched baseball
S = event that a viewer watched soccer

Then the event "the group that watched none of the three sports during the last year" is the set $(G \cup B \cup S)^c$
2.4 The events $R_1 \cap R_2$ and $B_1 \cap B_2$ represent the events that both ball are the same color and therefore as sets they are disjoint
2.5 880
2.6 50%
2.7 5%
2.8 60
2.9 53%

2.10 Using Theorem 2.3, we find

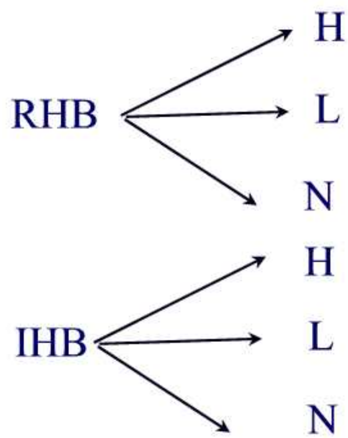
$$\begin{aligned}
n(A\cup B\cup C) =& n(A\cup(B\cup C))\\
=& n(A)+n(B\cup C)-n(A\cap(B\cup C))\\
=& n(A)+(n(B)+n(C)-n(B\cap C))\\
& -n((A\cap B)\cup(A\cap C))\\
=& n(A)+(n(B)+n(C)-n(B\cap C))\\
& -(n(A\cap B)+n(A\cap C)-n(A\cap B\cap C))\\
=& n(A)+n(B)+n(C)-n(A\cap B)-n(A\cap C)\\
& -n(B\cap C)+n(A\cap B\cap C)
\end{aligned}$$

2.11 50
2.12 10
2.13 (a) 3 (b) 6
2.14 20%
2.15 (a) Let $x\in A\cap(B\cup C)$. Then $x\in A$ and $x\in B\cup C$. Thus, $x\in A$ and ($x\in B$ or $x\in C$). This implies that ($x\in A$ and $x\in B$) or ($x\in A$ and $x\in C$). Hence, $x\in A\cap B$ or $x\in A\cap C$, i.e. $x\in(A\cap B)\cup(A\cap C)$. The converse is similar.
(b) Let $x\in A\cup(B\cap C)$. Then $x\in A$ or $x\in B\cap C$. Thus, $x\in A$ or ($x\in B$ and $x\in C$). This implies that ($x\in A$ or $x\in B$) and ($x\in A$ or $x\in C$). Hence, $x\in A\cup B$ and $x\in A\cup C$, i.e. $x\in(A\cup B)\cap(A\cup C)$. The converse is similar.
2.16 (a) $B\subseteq A$
(b) $A\cap B=\emptyset$ or $A\subseteq B^c$.
(c) $A\cup B-A\cap B$
(d) $(A\cup B)^c$
2.17 37

Section 3

3.1 (a) 100 (b) 900 (c) 5,040 (d) 90,000
3.2 (a) 336 (b) 6
3.3 6
3.4 90

3.5



3.6 4^{12}
3.7 380
3.8 255,024
3.9 5,040
3.10 384

Section 4

4.1 $m = 9$ and $n = 3$
4.2 (a) 456,976 (b) 358,800
4.3 (a) 15,600,000 (b) 11,232,000
4.4 (a) 64,000 (b) 59,280
4.5 (a) 479,001,600 (b) 604,800
4.6 (a) 5 (b) 20 (c) 60 (d) 120
4.7 60
4.8 15,600

Section 5

5.1 $m = 13$ and $n = 1$ or $n = 12$
5.2 11,480
5.3 300
5.4 10
5.5 28

5.6 4,060
5.7 Recall that ${}_mP_n = \frac{m!}{(m-n)!} = n!_mC_n$. Since $n! \geq 1$, we can multiply both sides by ${}_mC_n$ to obtain ${}_mP_n = n!_mC_n \geq_m C_n$.
5.8 (a) Combination (b) Permutation
5.9 $(a+b)^7 = a^7 + 7a^6b + 21a^5b^2 + 35a^4b^3 + 35a^3b^4 + 21a^2b^5 + 7ab^6 + b^7$
5.10 $22,680a^3b^4$
5.11 1,200

Section 6

6.1 (a) $S = \{1, 2, 3, 4, 5, 6\}$ (b) $\{2, 4, 6\}$
6.2 $S = \{(H, 1), (H, 2), (H, 3), (H, 4), (H, 5), (H, 6), (T, H), (T, T)\}$
6.3 50%
6.4 (a) $(i, j), i, j = 1, \cdots, 6$ (b) $E^c = \{(5, 6), (6, 5), (6, 6)\}$ (c) $\frac{11}{12}$ (d) $\frac{5}{6}$ (e) $\frac{7}{9}$
6.5 (a) 0.5 (b) 0 (c) 1 (d) 0.4 (e) 0.3
6.6 (a) 0.75 (b) 0.25 (c) 0.5 (d) 0 (e) 0.375 (f) 0.125
6.7 25%
6.8 $\frac{6}{128}$
6.9 $1 - 6.6 \times 10^{-14}$
6.10 (a) 10 (b) 40%
6.11 (a) $S = \{D_1D_2, D_1N_1, D_1N_2, D_1N_3, D_2N_1, D_2N_2, D_2N_3, N_1N_2, N_1N_3, N_2N_3\}$ (b) 10%

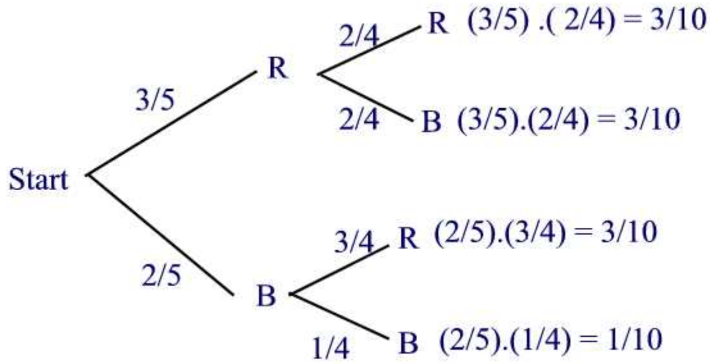
Section 7

7.1 (a) 0.78 (b) 0.57 (c) 0
7.2 0.32
7.3 0.308
7.4 0.555
7.5 Since $Pr(A \cup B) \leq 1$, we have $-Pr(A \cup B) \geq -1$. Add $Pr(A) + Pr(B)$ to both sides to obtain $Pr(A) + Pr(B) - Pr(A \cup B) \geq Pr(A) + Pr(B) - 1$. But the left hand side is just $Pr(A \cap B)$.
7.6 (a) 0.181 (b) 0.818 (c) 0.545
7.7 0.889
7.8 No
7.9 0.52
7.10 0.05
7.11 0.6

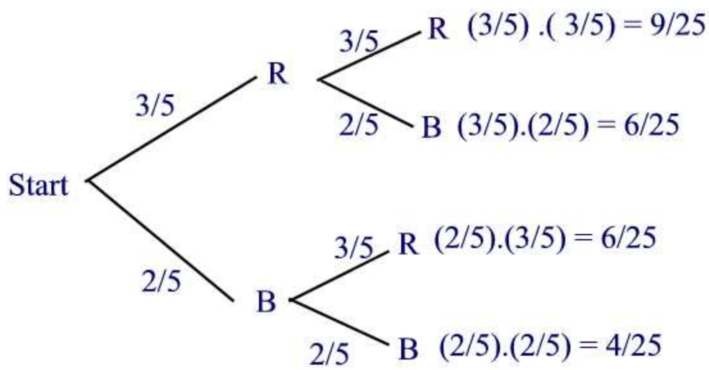
7.12 0.48
7.13 0.04
7.14 0.5
7.15 10%
7.16 80%
7.17 0.89

Section 8

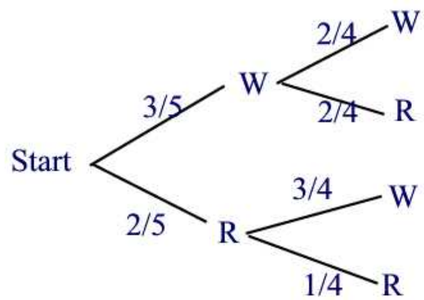
8.1



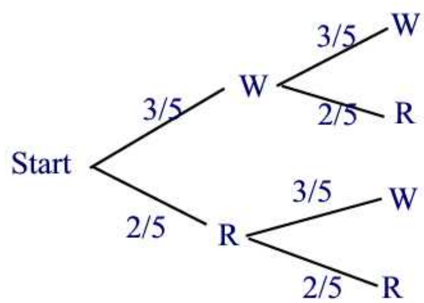
8.2



8.3 $\Pr(A) = 0.6, \Pr(B) = 0.3, \Pr(C) = 0.1$
8.4 0.1875
8.5 0.444
8.6 0.167
8.7 The probability is $\frac{3}{5} \cdot \frac{2}{4} + \frac{2}{5} \cdot \frac{3}{4} = \frac{3}{5} = 0.6$



8.8 The probability is $\frac{3}{5} \cdot \frac{2}{5} + \frac{2}{5} \cdot \frac{3}{5} = \frac{12}{25} = 0.48$



8.9 0.14
8.10 $\frac{36}{65}$
8.11 0.102
8.12 0.27

Section 9

9.1 0.173
9.2 0.205
9.3 0.467
9.4 0.5
9.5 (a) 0.19 (b) 0.60 (c) 0.31 (d) 0.317 (e) 0.613
9.6 0.151
9.7 0.133
9.8 0.978
9.9 $\frac{7}{1912}$
9.10 (a) $\frac{1}{221}$ (b) $\frac{1}{169}$
9.11 $\frac{1}{114}$
9.12 80.2%
9.13 (a) 0.021 (b) 0.2381, 0.2857, 0.476

9.14 (a) 0.57 (b) 0.211 (c) 0.651

Section 10

10.1 (a) 0.26 (b) $\frac{6}{13}$
10.2 0.1584
10.3 0.0141
10.4 0.29
10.5 0.42
10.6 0.22
10.7 0.657
10.8 0.4
10.9 0.45
10.10 0.66
10.11 (a) 0.22 (b) $\frac{15}{22}$
10.12 (a) 0.56 (b) $\frac{4}{7}$
10.13 0.36
10.14 $\frac{1}{3}$
10.15 (a) $\frac{17}{140}$ (b) $\frac{7}{17}$
10.16 $\frac{15}{74}$.
10.17 $\frac{72}{73}$.

Section 11

11.1 (a) Dependent (b) Independent
11.2 0.02
11.3 (a) 21.3% (b) 21.7%
11.4 0.72
11.5 4
11.6 0.328
11.7 0.4
11.8 We have

$$\begin{aligned}
\Pr(A\cap B) =&\Pr(\{1\}) = \frac{1}{4} = \frac{1}{2}\times\frac{1}{2} = \Pr(A)\Pr(B)\\
\Pr(A\cap C) =&\Pr(\{1\}) = \frac{1}{4} = \frac{1}{2}\times\frac{1}{2} = \Pr(A)\Pr(C)\\
\Pr(B\cap C) =&\Pr(\{1\}) = \frac{1}{4} = \frac{1}{2}\times\frac{1}{2} = \Pr(B)\Pr(C)
\end{aligned}$$

It follows that the events A, B, and C are pairwise independent. However,

$$\Pr(A \cap B \cap C) = \Pr(\{1\}) = \frac{1}{4} \neq \frac{1}{8} = \Pr(A)\Pr(B)\Pr(C).$$

Thus, the events A, B, and C are not independent
11.9 $\Pr(C) = 0.56,\ \Pr(D) = 0.38$
11.10 0.93
11.11 0.43
11.12 (a) We have $\Pr(A|B \cap C) = \frac{\Pr(A\cap B\cap C)}{\Pr(B\cap C)} = \frac{\Pr(A)\Pr(B)\Pr(C)}{\Pr(B)\Pr(C)} = \Pr(A)$. Thus, A and $B \cap C$ are independent.
(b) We have $\Pr(A|B\cup C) = \frac{\Pr(A\cap(B\cup C))}{\Pr(B\cup C)} = \frac{\Pr((A\cap B)\cup(A\cap C))}{\Pr(B\cup C)} = \frac{\Pr(A\cap B)+\Pr(A\cap C)-\Pr(A\cap B\cap C)}{\Pr(B)+\Pr(C)-\Pr(B\cap C)} = \frac{\Pr(A)\Pr(B)+\Pr(A)\Pr(C)-\Pr(A)\Pr(B)\Pr(C)}{\Pr(A)+\Pr(B)-\Pr(A)\Pr(B)} = \frac{\Pr(A)\Pr(B)[1-\Pr(C)]+\Pr(A)\Pr(C)}{\Pr(B)+\Pr(C)-\Pr(B)\Pr(C)} = \frac{\Pr(A)\Pr(B)\Pr(C^c)+\Pr(A)\Pr(C)}{\Pr(B)\Pr(C^c)+\Pr(C)} = \frac{\Pr(A)[\Pr(B)\Pr(C^c)+\Pr(C)]}{\Pr(B)\Pr(C^c)+\Pr(C)} = \Pr(A)$. Hence, A and $B \cup C$ are independent
11.13 (a) We have

$$\begin{aligned}
S =& \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}\\
A =& \{HHH, HHT, HTH, THH, TTH\}\\
B =& \{HHH, THH, HTH, TTH\}\\
C =& \{HHH, HTH, THT, TTT\}
\end{aligned}$$

(b) $\Pr(A) = \frac{5}{8},\ \Pr(B) = 0.5,\ \Pr(C) = \frac{1}{2}$
(c) $\frac{4}{5}$
(d) We have $B \cap C = \{HHH, HTH\}$, so $\Pr(B \cap C) = \frac{1}{4}$. That is equal to $\Pr(B)\Pr(C)$, so B and C are independent
11.14 0.65
11.15 (a) 0.70 (b) 0.06 (c) 0.24 (d) 0.72 (e) 0.4615

Section 12

12.1 15:1
12.2 62.5%
12.3 1:1
12.4 4:6
12.5 4%
12.6 (a) 1:5 (b) 1:1 (c) 1:0 (d) 0:1
12.7 1:3
12.8 (a) 43% (b) 0.3

Section 13

13.1 (a) Continuous (b) Discrete (c) Discrete (d) Continuous (e) mixed.
13.2 If G and R stand for golden and red, the probabilities for GG, GR, RG, and RR are, respectively $\frac{5}{8}\cdot\frac{4}{7}=\frac{5}{14}, \frac{5}{8}\cdot\frac{3}{7}=\frac{15}{56}, \frac{3}{8}\cdot\frac{5}{7}=\frac{15}{56}$, and $\frac{3}{8}\cdot\frac{2}{7}=\frac{3}{28}$. The results are shown in the following table.

Element of sample space	Probability	x
GG	$\frac{5}{14}$	2
GR	$\frac{15}{56}$	1
RG	$\frac{15}{56}$	1
RR	$\frac{3}{28}$	0

13.3 0.139
13.4 0.85
13.5 $\left(\frac{1}{2}\right)^n$
13.6 $\frac{1}{2}$
13.7 0.4
13.8 0.9722
13.9 (a)

$$X(s)=\begin{cases} 0 & s\in\{(NS,NS,NS)\} \\ 1 & s\in\{(S,NS,NS),(NS,S,NS),(NS,NS,S)\} \\ 2 & s\in\{(S,S,NS),(S,NS,S),(NS,S,S)\} \\ 3 & s\in\{(S,S,S)\} \end{cases}$$

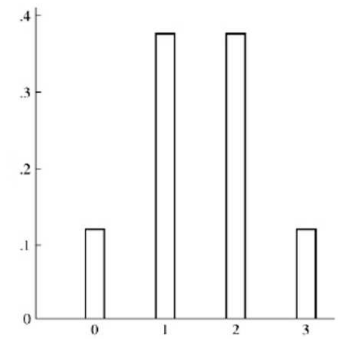
(b) 0.09 (c) 0.36 (d) 0.41 (e) 0.14
13.10 $\frac{1}{1+e}$
13.11 0.267

Section 14

14.1 (a)

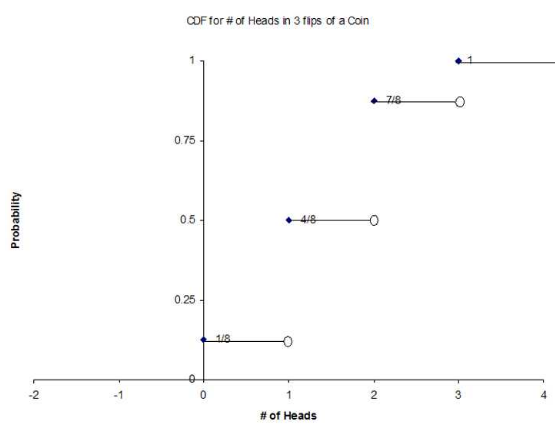
x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

(b)



14.2

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \le x < 1 \\ \frac{1}{2}, & 1 \le x < 2 \\ \frac{7}{8}, & 2 \le x < 3 \\ 1, & 3 \le x \end{cases}$$

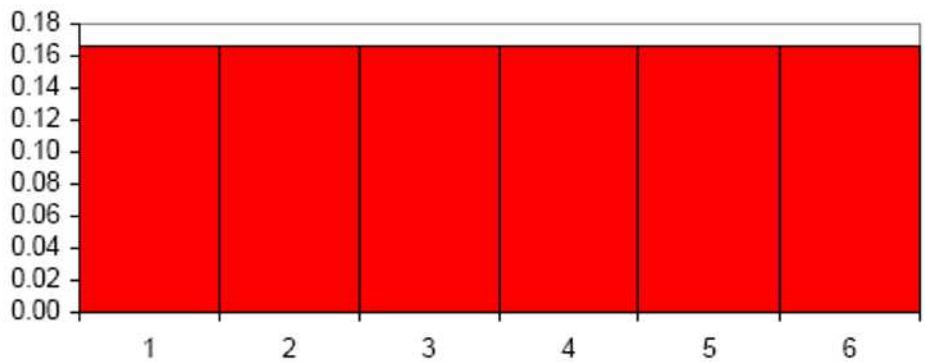


14.3

x	2	3	4	5	6	7	8	9	10	11	12
$p(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

14.4

X	Event	Pr(x)
1	R	$\frac{1}{6}$
2	(NR,R)	$\frac{5}{6} \cdot \frac{1}{5} = \frac{1}{6}$
3	(NR,NR,R)	$\frac{5}{6} \cdot \frac{4}{5} \cdot \frac{1}{4} = \frac{1}{6}$
4	(NR,NR,NR,R)	$\frac{5}{6} \cdot \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{6}$
5	(NR,NR,NR,NR,R)	$\frac{5}{6} \cdot \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{6}$
6	(NR,NR,NR,NR,NR,R)	$\frac{5}{6} \cdot \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{6}$



14.5

x	0	1	2
$p(x)$	1/4	1/2	1/4

14.6

$$\begin{aligned}F(n) =&\Pr(X \leq n) = \sum_{k=0}^{n} \Pr(X = k)\\=&\sum_{k=0}^{n} \frac{1}{3}\left(\frac{2}{3}\right)^k\\=&\frac{1}{3}\frac{1-\left(\frac{2}{3}\right)^{n+1}}{1-\frac{2}{3}}\\=&1-\left(\frac{2}{3}\right)^{n+1}\end{aligned}$$

14.7 (a) For $n = 2, 3, \cdots, 96$ we have

$$\Pr(X = n) = \frac{95}{100} \cdot \frac{94}{99} \cdots \frac{95 - n + 2}{100 - n + 2} \frac{5}{100 - n + 1}$$

and $\Pr(X=1) = \frac{5}{100}$
(b)

$$\Pr(Y=n) = \frac{\left(\begin{array}{c}5\\n\end{array}\right)\left(\begin{array}{c}95\\10-n\end{array}\right)}{\left(\begin{array}{c}100\\10\end{array}\right)}, \quad n = 0,1,2,3,4,5$$

14.8

$$p(x) = \begin{cases} \frac{3}{10} & x = -4 \\ \frac{4}{10} & x = 1 \\ \frac{3}{10} & x = 4 \end{cases}$$

14.9

$$\begin{aligned}\Pr(X=2) =&\Pr(RR) + \Pr(BB) = \frac{{}_3C_2}{{}_7C_2} + \frac{{}_4C_2}{{}_7C_2}\\ =&\frac{3}{21} + \frac{6}{21} = \frac{9}{21} = \frac{3}{7}\end{aligned}$$

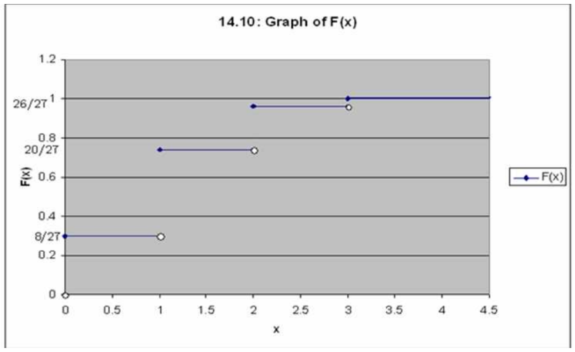
and

$$\Pr(X=-1) = 1 - \Pr(X=2) = 1 - \frac{3}{7} = \frac{4}{7}$$

14.10 (a)

$$\begin{aligned}p(0) =& \left(\frac{2}{3}\right)^3\\ p(1) =&3\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2\\ p(2) =&3\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)\\ p(3) =& \left(\frac{1}{3}\right)^3\end{aligned}$$

(b)



14.11

$$p(0) = \frac{220}{455}$$
$$p(1) = \frac{198}{455}$$
$$p(2) = \frac{36}{455}$$
$$p(3) = \frac{1}{455}$$

14.12 $p(2) = \frac{1}{36}, p(3) = \frac{2}{36}, p(4) = \frac{3}{36}, p(5) = \frac{4}{36}, p(6) = \frac{5}{36}, p(7) = \frac{6}{36}, p(8) = \frac{5}{36}, p(9) = \frac{4}{36}, p(10) = \frac{3}{36}, p(11) = \frac{2}{36}$, and $p(12) = \frac{1}{36}$ and 0 otherwise

Section 15

15.1 7
15.2 \$ 16.67
15.3 $E(X) = 10 \times \frac{1}{6} - 2 \times \frac{5}{6} = 0$ Therefore, you should come out about even if you play for a long time
15.4 -1
15.5 \$26
15.6 $E(X) = -\$0.125$ So the owner will make on average 12.5 cents per spin
15.7 \$110
15.8 -0.54
15.9 897
15.10 (a) 0.267 (b) 0.449 (c) 1.067

15.11 (a) 0.3 (b) 0.7 (c) $p(0) = 0.3,\ p(1) = 0,\ p(\Pr(X \leq 0)) = 0$ (d)

$$\Pr(x) = \begin{cases} 0.2 & x = -2 \\ 0.3 & x = 0 \\ 0.1 & x = 2.2 \\ 0.3 & x = 3 \\ 0.1 & x = 4 \\ 0 & \text{otherwise} \end{cases}$$

(e) 1.12
15.12 (a) 390 (b) Since $E(V) < 400$ the answer is no.
15.13 (a)

$$\begin{aligned} p(1) =& \Pr(X = 1) = \frac{{}_3C_3 \cdot {}_7C_1}{210} = \frac{7}{210} \\ p(2) =& \Pr(X = 2) = \frac{{}_3C_2 \cdot {}_7C_2}{210} = \frac{63}{210} \\ p(3) =& \Pr(X = 3) = \frac{{}_3C_1 \cdot {}_7C_3}{210} = \frac{105}{210} \\ p(4) =& \Pr(X = 4) = \frac{{}_3C_0 \cdot {}_7C_4}{210} = \frac{35}{210} \end{aligned}$$

(b)

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{7}{210} & 1 \leq x < 2 \\ \frac{70}{210} & 2 \leq x < 3 \\ \frac{175}{210} & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

(c) 2.8
15.14 \$50,400

Section 16

16.1 (a) $c = \frac{1}{30}$ (b) 3.333 (c) 8.467
16.2 (a) $c = \frac{1}{9}$ (b) $p(-1) = \frac{2}{9}, p(1) = \frac{3}{9}, p(2) = \frac{4}{9}$ (c) $E(X) = 1$ and $E(X^2) = \frac{7}{3}$
16.3 (a)

$$p(x) = \begin{cases} \frac{x}{21} & x = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise} \end{cases}$$

(b) $\frac{4}{7}$ (c) $E(X) = 4.333$
16.4 Let D denote the range of X. Then

$$\begin{aligned}E(aX^2 + bX + c) =& \sum_{x \in D}(ax^2 + bx + c)p(x)\\ =& \sum_{x\in D} ax^2p(x) + \sum_{x\in D} bxp(x) + \sum_{x\in D} cp(x)\\ =& a\sum_{x\in D} x^2p(x) + b\sum_{x\in D} xp(x) + c\sum_{x\in D} p(x)\\ =& aE(X^2) + bE(X) + c\end{aligned}$$

16.5 0.62
16.6 \$220
16.7 0.0314
16.8 0.24
16.9 (a)

$$\begin{aligned}\Pr(X = 2) =& \Pr(RR) + \Pr(BB) = \frac{{}_3C_2}{{}_7C_2} + \frac{{}_4C_2}{{}_7C_2}\\ =& \frac{3}{21} + \frac{6}{21} = \frac{9}{21} = \frac{3}{7}\end{aligned}$$

and

$$\Pr(X = -1) = 1 - \Pr(X = 2) = 1 - \frac{3}{7} = \frac{4}{7}$$

(b) $E(2^X) = 2$
16.10 (a) $P = 3C + 8A + 5S - 300$ (b) \$1,101

Section 17

17.1 0.45
17.2 374
17.3 (a) $c = \frac{1}{55}$ (b) $E(X) = -1.09$ $\text{Var}(X) = 1.064$

17.4 (a)

$$\begin{aligned}
p(1) =& \frac{{}_3C_3 \cdot {}_7C_1}{{}_{10}C_4} = \frac{1}{30} \\
p(2) =& \frac{{}_3C_2 \cdot {}_7C_2}{{}_{10}C_4} = \frac{3}{10} \\
p(3) =& \frac{{}_3C_1 \cdot {}_7C_3}{{}_{10}C_4} = \frac{1}{2} \\
p(4) =& \frac{{}_3C_0 \cdot {}_7C_4}{{}_{10}C_4} = \frac{1}{6}
\end{aligned}$$

(b) $E(X) = 2.8$ $\text{Var}(X) = 0.56$
17.5 (a) $c = \frac{1}{30}$ (b) $E(X) = 3.33$ and $E(X(X-1)) = 8.467$ (c) $E(X^2) = 11.8$ and $\text{Var}(X) = 0.6889$
17.6 $E(Y) = 21$ and $\text{Var}(Y) = 144$
17.7 (a)

$$\begin{aligned}
\Pr(X = 2) =& \Pr(RR) + \Pr(BB) = \frac{{}_3C_2}{{}_7C_2} + \frac{{}_4C_2}{{}_7C_2} \\
=& \frac{3}{21} + \frac{6}{21} = \frac{9}{21} = \frac{3}{7}
\end{aligned}$$

and

$$\Pr(X = -1) = 1 - \Pr(X = 2) = 1 - \frac{3}{7} = \frac{4}{7}$$

(b) $E(X) = \frac{2}{7}$ and $E(X^2) = \frac{16}{7}$ (c) $\text{Var}(X) = \frac{108}{49}$
17.8 $E(X) = \frac{4}{10}$; $\text{Var}(X) = 9.84$; $\sigma_X = 3.137$
17.9 $E(X) = p$ $\text{Var}(X) == p(1-p)$

Section 18

18.1 0.3826
18.2 0.211
18.3

$$p(x) = \begin{cases} \frac{1}{8}, & \text{if } x = 0, 3 \\ \frac{3}{8}, & \text{if } x = 1, 2 \\ 0, & \text{otherwise} \end{cases}$$

18.4 0.096
18.5 \$60

18.6 0.925
18.7 0.144
18.8 (a) 0.5 (b) 0.1875
18.9 0.1198
18.10 0.6242
18.11 0.784
18.12 (a) 0.2321 (b) 0.2649 (c) 0.1238
18.13 0.2639

Section 19

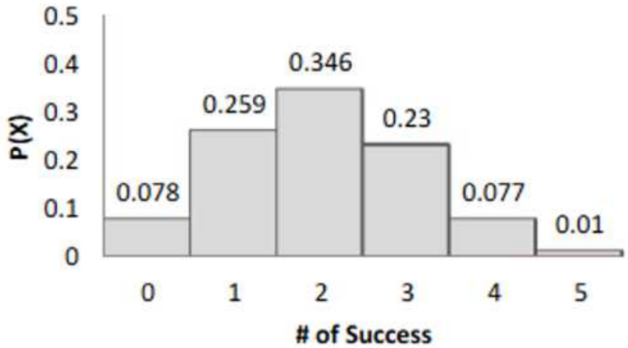
19.1 154
19.2 \$985
19.3

n	$E(Y)$	$E(Y^2)$	$E(S)$
1	0.20	0.20	100+10-2 = 108
2	0.40	0.48	100+20-4.8 = 115.2
3	0.60	0.84	100+30-8.4 = 121.6

19.4 $E(X) = 400$ and $\sigma_X = 15.492$
19.5 (a) $p(x) = {}_5C_x(0.4)^x(0.6)^{5-x}, \ x = 0, 1, 2, 3, 4, 5.$
(b)

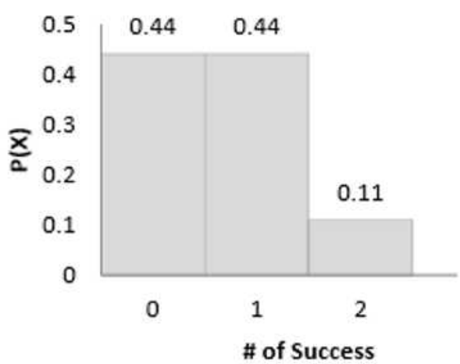
$$\begin{aligned}
p(0) =& {}_5C_0(0.4)^0(0.6)^{5-0} = 0.078\\
p(1) =& {}_5C_1(0.4)^1(0.6)^{5-1} = 0.259\\
p(2) =& {}_5C_2(0.4)^2(0.6)^{5-2} = 0.346\\
p(3) =& {}_5C_3(0.4)^3(0.6)^{5-3} = 0.230\\
p(4) =& {}_5C_4(0.4)^4(0.6)^{5-4} = 0.077\\
p(5) =& {}_5C_5(0.4)^5(0.6)^{5-5} = 0.01.
\end{aligned}$$

(c)



(d) $E(X) = 2$ and $\sigma_X = 1.095$.
19.6 (a) $p(x) = {}_2C_x \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{2-x}$.
(b)

$$p(0) = {}_2C_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^{2-0} = 0.44$$
$$p(1) = {}_2C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^{2-1} = 0.44$$
$$p(2) = {}_2C_2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^{2-2} = 0.11.$$



(c)

(d) $E(X) = 0.667$ and $\sigma_X = 0.667$.

Section 20

20.1 0.0183 and 0.0027
20.2 0.1251
20.3 3.06×10^{-7}
20.4 (a) 0.577 (b) 0.05
20.5 (a) 0.947 (b) 0.762 (c) 0.161
20.6 0.761897
20.7 (a) 0.5654 (b) 0.4963
20.8 2
20.9 $7231.30
20.10 699
20.11 0.1550
20.12 0.7586
20.13 4
20.14 (a) 0.2873 (b) mean = 20 and standard deviation = 4.47

Section 21

21.1 $n \geq 20$ and $p \leq 0.05$. $\Pr(X \geq 2) \approx 0.9084$
21.2 0.3293
21.3 0.0144
21.4 0.3679
21.5 (a) 0.177 (b) 0.876

Section 22

22.1 (a) 0.1 (b) 0.09 (c) $\left(\frac{9}{10}\right)^{n-1}\left(\frac{1}{10}\right)$
22.2 0.387
22.3 0.916
22.4 (a) 0.001999 (b) 1000
22.5 (a) 0.1406 (b) 0.3164
22.6 (a) 0.1481 and 7.842×10^{-10} (b) 3
22.7 (a) $p_X(x) = (0.85)(0.15)^{x-1}, \quad x = 1, 2, \cdots$ and 0 otherwise (b) $p_Y(y) = (0.85)(0.15)^{y-1}, \quad y = 1, 2, 3, \cdots$ and 0 otherwise. Thus, Y is a geometric random variable with parameter 0.85
22.8 (a) 0.1198 (b) 0.3999

22.9 We have

$$\begin{aligned}\Pr(X > i+j|X > i) =& \frac{\Pr(X > i+j, X > i)}{\Pr(X > i)}\\ =& \frac{\Pr(X > i+j)}{\Pr(X > i)} = \frac{(1-p)^{i+j}}{(1-p)^i}\\ =& (1-p)^j = \Pr(X > j)\end{aligned}$$

22.10 0.053
22.11 (a) 10 (b) 0.81
22.12 (a) X is a geometric distribution with pmf $p(x) = 0.4(0.6)^{x-1}, \; x = 1, 2, \cdots$ (b) X is a binomial random variable with pmf $p(x) = {}_{20}C_x(0.60)^x(0.40)^{20-x}$ where $x = 0, 1, \cdots, 20$

Section 23

23.1 (a) 0.0103 (b) $E(X) = 80; \sigma_X = 26.833$
23.2 0.0307
23.3 (a) X is negative binomial distribution with $r = 3$ and $p = \frac{4}{52} = \frac{1}{13}$. So $p(n) = {}_{k-1}C_2 \left(\frac{1}{13}\right)^3 \left(\frac{12}{13}\right)^{n-3}$ (b) 0.01793
23.4 $E(X) = 24$ and $\text{Var}(X) = 120$
23.5 0.109375
23.6 0.1875
23.7 0.2898
23.8 0.022
23.9 (a) 0.1198 (b) 0.0254
23.10 $E(X) = \frac{r}{p} = 20$ and $\sigma_X = \sqrt{\frac{r(1-p}{p^2}} = 13.416$
23.11 0.0645
23.12 ${}_{n-1}C_2 \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^{n-3}$
23.13 3

Section 24

24.1 0.32513
24.2 0.1988
24.3

k	0	1	2	3	4	5	6
$\Pr(X = k)$	0.468	0.401	0.117	0.014	7.06×10^{-4}	1.22×10^{-5}	4.36×10^{-8}

24.4 0.247678
24.5 0.073
24.6 (a) 0.214 (b) $E(X) = 3$ and $\text{Var}(X) = 0.429$
24.7 0.793
24.8 $\frac{{}_{2477}C_3 \cdot {}_{121373}C_{97}}{{}_{123850}C_{100}}$
24.9 0.033
24.10 0.2880
24.11 0.375
24.12 0.956

Section 25

25.1 (a)

$$\begin{aligned}
\Pr(2) =& \frac{C(2,2)}{C(5,2)} = 0.1 \\
\Pr(6) =& \frac{C(2,1)C(2,1)}{C(5,2)} = 0.4 \\
\Pr(10) =& \frac{C(2,2)}{C(5,2)} = 0.1 \\
\Pr(11) =& \frac{C(1,1)C(2,1)}{C(5,2)} = 0.2 \\
\Pr(15) =& \frac{C(1,1)C(2,1)}{C(5,2)} = 0.2
\end{aligned}$$

(b)

$$F(x) = \begin{cases} 0 & x < 2 \\ 0.1 & 2 \le x < 6 \\ 0.5 & 6 \le x < 10 \\ 0.6 & 10 \le x < 11 \\ 0.8 & 11 \le x < 15 \\ 1 & x \ge 15 \end{cases}$$

(c) 8.8
25.2

x	$(-\infty,0)$	$[0,1)$	$[1,2)$	$[2,3)$	$[3,\infty)$
$\Pr(X\le x)$	0	0.495	0.909	0.996	1

25.3 (a)

$$\begin{aligned}
\Pr(X=1)=&\Pr(X\le 1)-\Pr(X<1)=F(1)-\lim_{n\to\infty}F\left(1-\frac{1}{n}\right)\\
=&\frac{1}{2}-\frac{1}{4}=\frac{1}{4}\\
\Pr(X=2)=&\Pr(X\le 2)-\Pr(X<2)=F(2)-\lim_{n\to\infty}F\left(2-\frac{1}{n}\right)\\
=&\frac{11}{12}-\left(\frac{1}{2}-\frac{2-1}{4}\right)=\frac{1}{6}\\
\Pr(X=3)=&\Pr(X\le 3)-\Pr(X<3)=F(3)-\lim_{n\to\infty}F\left(3-\frac{1}{n}\right)\\
=&1-\frac{11}{12}=\frac{1}{12}
\end{aligned}$$

(b) 0.5

25.4

$$\begin{aligned}
\Pr(X=0)=&F(1)=\frac{1}{2}\\
\Pr(X=1)=&F(1)-F(1^-)=\frac{3}{5}-\frac{1}{2}=\frac{1}{10}\\
\Pr(X=2)=&F(2)-F(2^-)=\frac{4}{5}-\frac{3}{5}=\frac{1}{5}\\
\Pr(X=3)=&F(3)-F(3^-)=\frac{9}{10}-\frac{4}{5}=\frac{1}{10}\\
\Pr(X=3.5)=&F(3.5)-F(3.5^-)=1-\frac{9}{10}=\frac{1}{10}
\end{aligned}$$

and 0 otherwise.

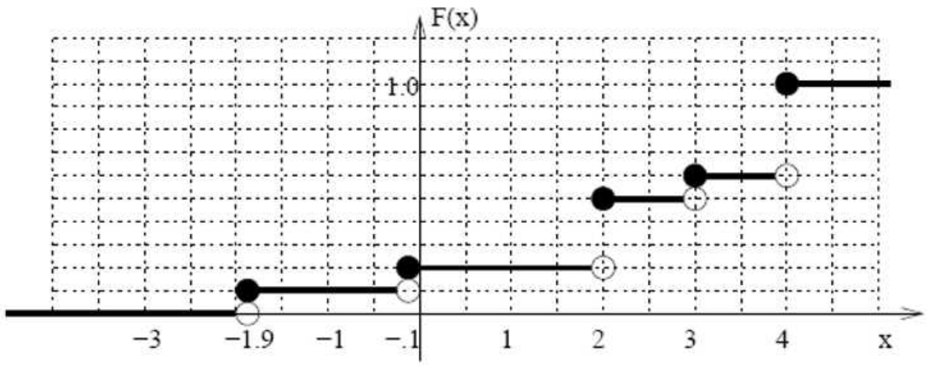
25.5 (a)

$$\Pr(x)=\begin{cases}0.1 & x=-2\\ 0.2 & x=1.1\\ 0.3 & x=2\\ 0.4 & x=3\\ 0 & \text{otherwise}\end{cases}$$

(b) 0 (c) 0.4 (d) 0.444
25.6 (a) 0.1
(b)

$$F(x) = \begin{cases} 0 & x < -1.9 \\ 0.1 & -1.9 \leq x < -0.1 \\ 0.2 & -0.1 \leq x < 2 \\ 0.5 & 2 \leq x < 3 \\ 0.6 & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

The graph of $F(x)$ is shown below.



(c) $F(0) = 0.2; F(2) = 0.5; F(F(3.1)) = 0.2.$ (d) 0.5 (e) 0.64
25.7 (a)

$$\Pr(x) = \begin{cases} 0.3 & x = -4 \\ 0.4 & x = 1 \\ 0.3 & x = 4 \\ 0 & \text{otherwise} \end{cases}$$

(b)$E(X) = 0.4,\ \text{Var}(X) = 9.84$, and $SD(X) = 3.137$
25.8 (a)

$$\Pr(x) = \begin{cases} \frac{1}{12} & x = 1, 3, 5, 8, 10, 12 \\ \frac{2}{12} & x = 2, 4, 6 \\ 0 & \text{otherwise} \end{cases}$$

(b)

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{12} & 1 \le x < 2 \\ \frac{3}{12} & 2 \le x < 3 \\ \frac{4}{12} & 3 \le x < 4 \\ \frac{6}{12} & 4 \le x < 5 \\ \frac{7}{12} & 5 \le x < 6 \\ \frac{9}{12} & 6 \le x < 8 \\ \frac{10}{12} & 8 \le x < 10 \\ \frac{11}{12} & 10 \le x < 12 \\ 1 & x \ge 12 \end{cases}$$

(c) $\Pr(X < 4) = 0.333$. This is not the same as $F(4)$ which is the probability that $X \le 4$. The difference between them is the probability that X is EQUAL to 4.

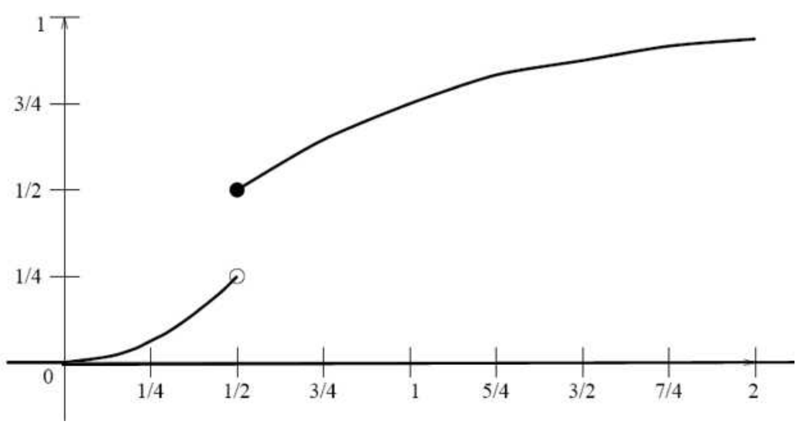
25.9 (a) We have

$$\begin{aligned}
\Pr(X = 0) =& \text{(jump in } F(x) \text{ at } x = 0) = 0 \\
\Pr(X = 1) =& \text{(jump in } F(x) \text{ at } x = 1) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4} \\
\Pr(X = 2) =& \text{(jump in } F(x) \text{ at } x = 2) = 1 - \frac{3}{4} = \frac{1}{4}
\end{aligned}$$

(b) $\frac{7}{16}$ (c) $\frac{3}{16}$ (d) $\frac{3}{8}$

25.10 (a) 0.125 (b) 0.584 (c) 0.5 (d) 0.25

(e)



Section 26

26.1 (a) We have

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - \frac{1}{10}(100 - x)^{\frac{1}{2}}, & 0 \le x \le 100 \\ 1, & x > 100. \end{cases}$$

(b) 0.092
26.2 (a) 0.3 (b) 0.3
26.3 Using the fundamental theorem of calculus, we have

$$S'(x) = \left(\int_x^\infty f(t)dt\right)' = \left(-\int_\infty^x f(t)dt\right)' = -f(x)$$

26.4

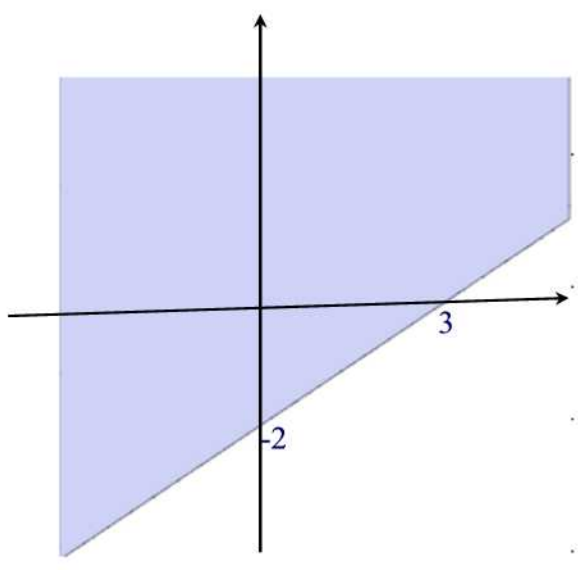
$$f(x) = \begin{cases} 0, & x \le 0 \\ \lambda e^{-\lambda x}, & x > 0. \end{cases}$$

26.5

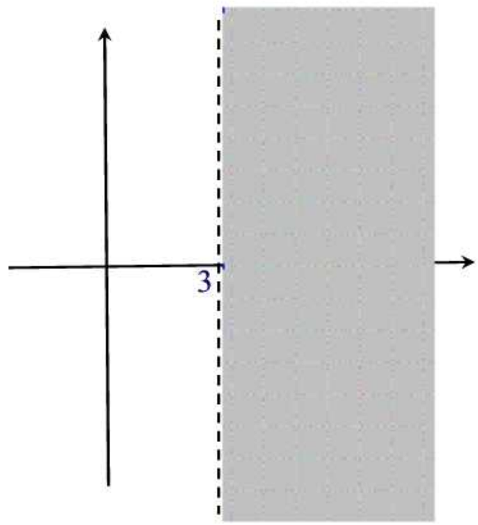
$$S(x) = \begin{cases} 1, & x \le 0 \\ 1 - x, & 0 < x < 1 \\ 0, & x \ge 1 \ \blacksquare \end{cases}$$

Section 27

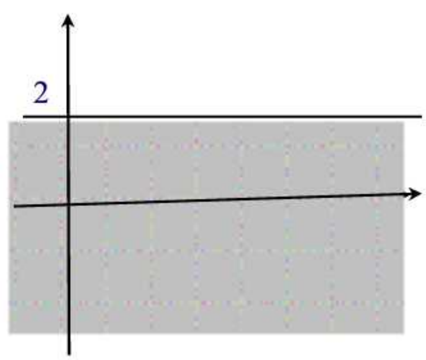
27.1



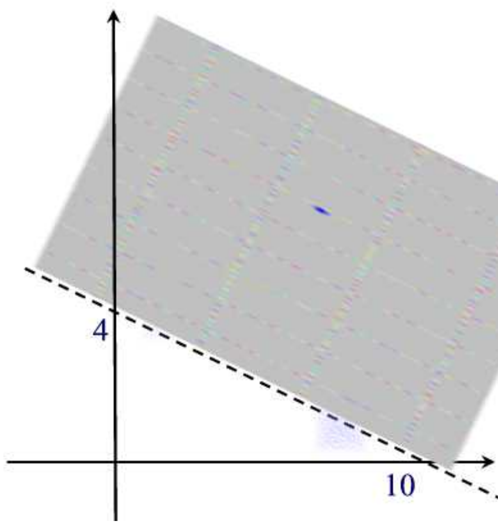
27.2



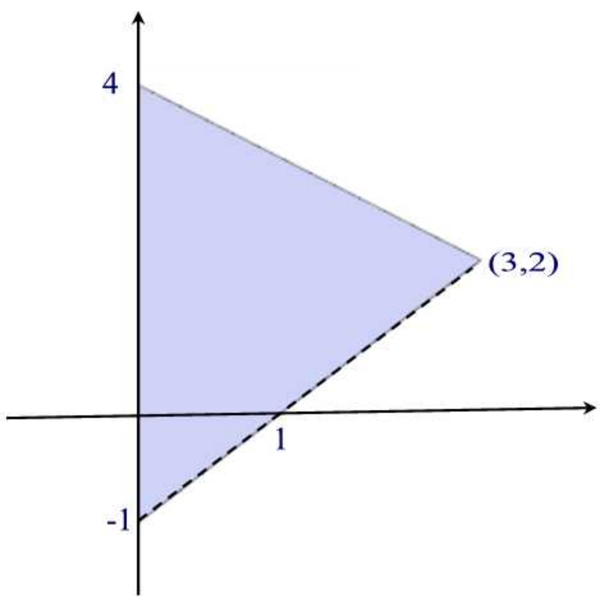
27.3



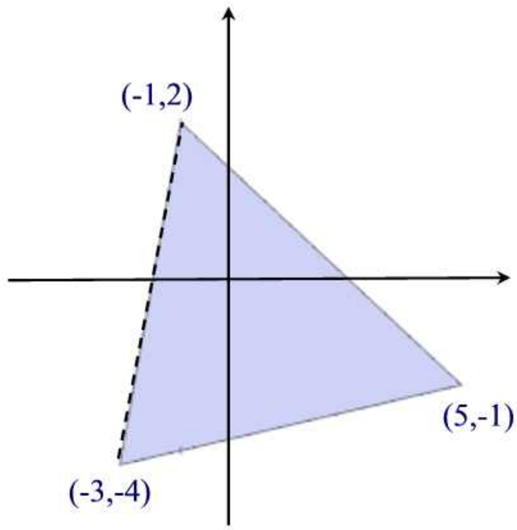
27.4



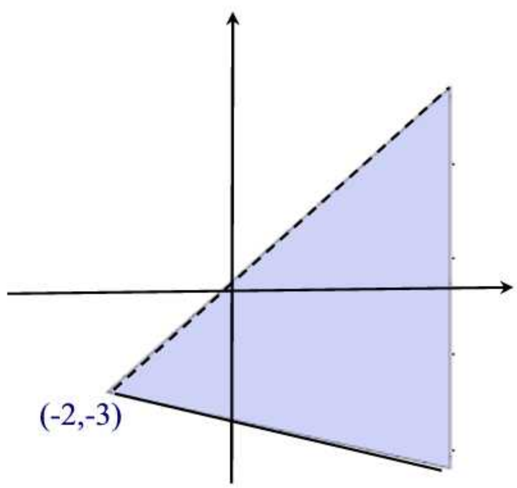
27.5



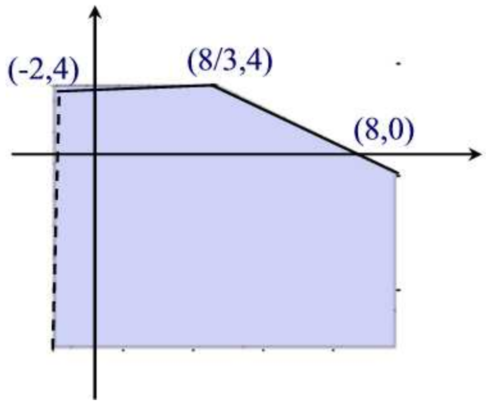
27.6



27.7



27.8



Section 28

28.1 Divergent
28.2 Divergent
28.3 Divergent
28.4 6
28.5 $\frac{\pi}{2}$
28.6 $\frac{\pi}{4}$
28.7 1
28.8 Divergent
28.9 Divergent
28.10 Divergent
28.11 2π
28.12 1
28.13 $\frac{2}{3}$
28.14 Divergent
28.15 Divergent
28.16 Divergent
28.17 convergent

Section 29

29.1 $\int_0^1 \int_x^{x+1} f(x,y)dydx$
29.2

$$\int_0^1 \int_0^{\frac{x}{2}} f(x,y)dydx \text{ or } \int_0^{\frac{1}{2}} \int_{2y}^1 f(x,y)dxdy$$

29.3 $\int_{\frac{1}{2}}^1 \int_{\frac{1}{2}}^1 f(x,y)dydx$
29.4 $\int_0^{\frac{1}{2}} \int_{\frac{1}{2}}^1 f(x,y)dydx + \int_{\frac{1}{2}}^1 \int_0^1 f(x,y)dydx$
29.5 $\int_{20}^{30} \int_{20}^{50-x} f(x,y)dydx$
29.6 $\int_0^1 \int_0^{x+1} f(x,y)dydx + \int_1^{\infty} \int_{x-1}^{x+1} f(x,y)dydx$
29.7 $1 - 2e^{-1}$
29.8 $\frac{1}{6}$
29.9 $\frac{L^4}{3}$
29.10 $\int_{0.5}^1 \int_{0.5}^1 f(x,y)dydx$
29.11 $\frac{7}{8}$
29.12 $\frac{3}{8}$

Section 30

30.1 2
30.2 (a) 0.135 (b) 0.233
(c)

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{5}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

30.3 $k = 0.003$, 0.027
30.4 0.938
30.5 (b)

$$f(x) = F'(x) = \begin{cases} 0 & x < 0 \\ 1/2 & 0 \leq x < 1 \\ 1/6 & 1 \leq x < 4 \\ 0 & x \geq 4 \end{cases}$$

30.6 (a) 1 (b) 0.736
30.7(a) $f(x) = F'(x) = \frac{e^x}{(1+e^x)^2}$ (b) 0.231
30.8 About 4 gallons
30.9 $\frac{1}{9}$
30.10 0.469
30.11 0.132
30.12 0.578
30.13 0.3
30.14 $\frac{511}{512}$

30.15 2

Section 31

31.1 (a) 1.2
(b) The cdf is given by

$$F(x) = \begin{cases} 0 & x \le -1 \\ 0.2 + 0.2x & -1 < x \le 0 \\ 0.2 + 0.2x + 0.6x^2 & 0 < x \le 1 \\ 1 & x > 1 \end{cases}$$

(c) 0.25 (d) 0.4
31.2 (a) $a = \frac{3}{5}$ and $b = \frac{6}{5}$.
(b)

$$F(x) = \int_{-\infty}^{x} f(u)du = \begin{cases} \int_{-\infty}^{x} 0du = 0 & x < 0 \\ \int_{-\infty}^{x} \frac{1}{5}(3 + 5u^2)du = \frac{3}{5}x + \frac{2}{5}x^3 & 0 \le x \le 1 \\ \int_0^1 \frac{1}{5}(3 + 6u^2)du = 1 & x > 1 \end{cases}$$

31.3 (a) 4 (b) 0 (c) ∞
31.4 $E(X) = \frac{2}{3}$ and $\text{Var}(X) = \frac{2}{9}$
31.5 (a) $E(X) = \frac{1}{3}$ and $\text{Var}(X) = \frac{2}{9}$. $SD =\approx 0.471$ (b) 0.938
31.6 0.5
31.7 .06
31.8 $E(Y) = \frac{7}{3}$ and $\text{Var}(Y) = 0.756$
31.9 1.867
31.10 \$328
31.11 0.5
31.12 0.139
31.13 2.227
31.14 2694
31.15 6
31.16 347.96

Section 32

32.1 The median is 1 and the 70^{th} percentile is 2
32.2 0.693
32.3 $a + 2\sqrt{\ln 2}$

32.4 $3\ln 2$
32.5 0.8409
32.6 3659
32.7 1120 is the twentieth percentile or 1120 is the one-fifth quantile
32.8 3
32.9 2
32.10 $-\ln(1-p)$
32.11 $-\ln[2(1-p)]$
32.12 2
32.13 0.4472
32.14 6299.61
32.15 2.3811
32.16 2.71

Section 33

33.1 (a) The pdf is given by

$$f(x) = \begin{cases} \frac{1}{4} & 3 \le x \le 7 \\ 0 & \text{otherwise} \end{cases}$$

(b) 0 (c) 0.5
33.2 (a) The pdf is given by

$$f(x) = \begin{cases} \frac{1}{10} & 5 \le x \le 15 \\ 0 & \text{otherwise} \end{cases}$$

(b) 0.3 (c) $E(X) = 10$ and $\text{Var}(X) = 8.33$
33.3 (a)

$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

(b) $\Pr(a \le X \le a+b) = F(a+b) - F(a) = a + b - a = b$
33.4 $\frac{1}{n+1}$
33.5 0.693
33.6 0.667
33.7 500
33.8 1.707

33.9 403.44
33.10 (a) $\frac{e^2-1}{2e}$ (b) $\frac{1}{2}-\frac{1}{2e^2}$
33.11 3
33.12 $\frac{7}{10}$

Section 34

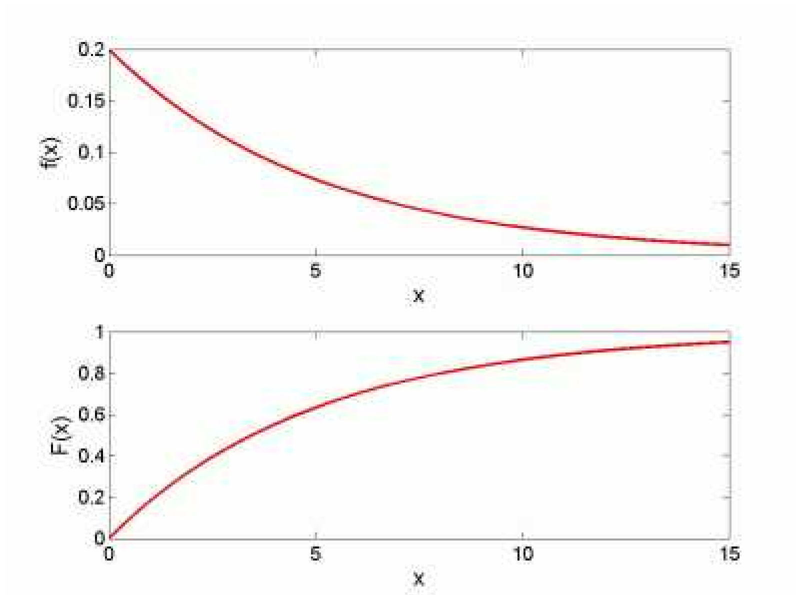
34.1 (a) 0.2389 (b) 0.1423 (c) 0.6188 (d) 88
34.2 (a) 0.7517 (b) 0.8926 (c) 0.0238
34.3 (a) 0.5 (b) 0.9876
34.4 (a) 0.4772 (b) 0.004
34.5 0.0228
34.6 (a) 0.9452 (b) 0.8186
34.7 0.223584
34.8 75
34.9 0.4721 (b) 0.1389 (c) 0.6664 (d) 0.58
34.10 0.86
34.11 (a) 0.1056 (b) 362.84
34.12 1:18pm

Section 35

35.1 (a) 0.2578 (b) 0.832
35.2 0.1788
35.3 (a) 0.8281 (b) 0.0021
35.4 0.0158
35.5 0.9854

Section 36

36.1 0.593
36.2



36.3 0.393
36.4 0.1175
36.5 0.549
36.6 (a) 0.189 (b) 0.250
36.7 (a) 0.1353 (b) 0.167 (c) 0.05
36.8 0.134
36.9 0.186
36.10 0.435
36.11 10256
36.12 3.540
36.13 5644.23
36.14 173.29
36.15 0.420
36.16 $\frac{4}{(\ln 2)^2}$
36.17 0.727

Section 37

37.1 We have

$$\begin{aligned} F_Y(y) =&\Pr(Y \le y) = \Pr\left(X \le \frac{y}{c}\right) \\ =&\frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^{\frac{y}{c}} t^{\alpha-1}e^{-\lambda t}dt \\ =&\frac{(\lambda/c)^\alpha}{\Gamma(\alpha)}\int_0^y z^{\alpha-1}e^{-\lambda\frac{z}{c}}dz. \end{aligned}$$

37.2 $E(X) = 1.5$ and $\mathrm{Var}(X) = 0.75$
37.3 0.0948
37.4 0.014
37.5 480
37.6 For $t \ge 0$ we have

$$F_{X^2}(t) = \Pr(X^2 \le t) = \Pr(-\sqrt{t} < X < \sqrt{t}) = \Phi(\sqrt{t}) - \Phi(-\sqrt{t})$$

Now, taking the derivative (and using the chain rule) we find

$$\begin{aligned} f_{X^2}(t) =&\frac{1}{2\sqrt{t}}\Phi'(\sqrt{t}) + \frac{1}{2\sqrt{t}}\Phi'(-\sqrt{t}) \\ =&\frac{1}{\sqrt{t}}\Phi'(\sqrt{t}) = \frac{1}{\sqrt{2\pi}}t^{-\frac{1}{2}e^{-\frac{1}{2}}} \end{aligned}$$

which is the density function of gamma distribution with $\alpha = \lambda = \frac{1}{2}$
37.7 $E(e^{tX}) = \left(\frac{\lambda}{\lambda-t}\right)^\alpha,\ t < \lambda$
37.8 We have

$$f'(x) = -\frac{\lambda^2 e^{-\lambda x}(\lambda x)^{\alpha-2}}{\Gamma(\alpha)}(\lambda x - \alpha + 1).$$

Thus, the only critical point of $f(x)$ is $x = \frac{1}{\lambda}(\alpha - 1)$. One can easily show that $f''\left(\frac{1}{\lambda}(\alpha-1)\right) < 0$
37.9 (a) The density function is

$$f(x) = \begin{cases} \frac{x^2}{432}e^{-\frac{x}{6}} & x \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

$E(X) = 18$ and $\sigma = 10.39$ (b) 1313
37.10 The density function is

$$f(x) = \begin{cases} \frac{2^{-\frac{n}{2}}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}}{\Gamma\left(\frac{n}{2}\right)} & x \ge 0 \\ 0 & \text{elsewhere} \end{cases}$$

The expected value is $E(X) = n$ and the variance is $\text{Var}(X) = 2n$.

Section 38

38.1 $f_Y(y) = \frac{1}{|a|\sqrt{2\pi}} e^{-\frac{(\frac{y-b}{a}-\mu)^2}{2}},\ a \neq 0$
38.2 $f_Y(y) = \frac{2(y+1)}{9}$ for $-1 \leq y \leq 2$ and $f_Y(y) = 0$ otherwise
38.3 For $0 \leq y \leq 8$ we have $f_Y(y) = \frac{y^{-\frac{1}{3}}}{6}$ and 0 otherwise
38.4 For $y \geq 1$ we have $f_Y(y) = \lambda y^{-\lambda-1}$ and 0 otherwise
38.5 For $y > 0$ we have $f_Y(y) = \frac{c}{m}\sqrt{\frac{2y}{m}} e^{-\frac{2\beta y}{m}}$ and 0 otherwise
38.6 For $y > 0$ we have $f_Y(y) = e^{-y}$ and 0 otherwise
38.7 For $-1 < y < 1$ we have $f_Y(y) = \frac{1}{\pi}\frac{1}{\sqrt{1-y^2}}$ and 0 otherwise
38.8 (a) For $0 < y < 1$ we have $f_Y(y) = \frac{1}{\alpha} y^{\frac{1-\alpha}{\alpha}}$ and 0 otherwise, $E(Y) = \frac{1}{\alpha+1}$
(b) For $y < 0$ we have $f_Y(y) = e^y$ and 0 otherwise, $E(Y) = -1$
(c) For $1 < y < e$ we have $f_Y(y) = \frac{1}{y}$ and 0 otherwise, $E(Y) = \int_1^e dy = e - 1$.
(d) For $0 < y < 1$ we have $f_Y(y) = \frac{2}{\pi\sqrt{1-y^2}}$, $E(Y) = \frac{2}{\pi}$ and 0 otherwise.
38.9 For $y > 4$ $f_Y(y) = 4y^{-2}$ and 0 otherwise
38.10 For $10,000e^{0.04} < v < 10,000e^{0.08}$ $F_V(v) = F_R(g^{-1}(v)) = 25\left(\ln\left(\frac{v}{10,000}\right) - 0.04\right)$ and $F_V(v) = 0$ for $v \leq 10,000e^{0.04}$ and $F_V(v) = 1$ for $v \geq 10,000e^{0.08}$
38.11 For $y > 0$ we have $f_Y(y) = \frac{1}{8}\left(\frac{y}{10}\right)^{\frac{1}{4}} e^{-\left(\frac{y}{10}\right)^{\frac{5}{4}}}$ and 0 otherwise
38.12 $f_R(r) = \frac{5}{2r^2}$ for $\frac{5}{6} < r < \frac{5}{4}$ and 0 otherwise
38.13 $f_Y(y) = \frac{1}{2} f_X\left(\frac{y}{2}\right)$ where X and Y are the monthly profits of Company A and Company B, respectively.
38.14 (a) 0.341 (b) $f_Y(y) = \frac{1}{2y\sqrt{2\pi}} exp\left(-\frac{1}{2 \cdot 2^2}(\ln y - 1)^2\right)$ for $y > 0$ and 0 otherwise.
38.15 The cdf is given by

$$F_Y(a) = \begin{cases} 0 & a \leq 0 \\ \sqrt{a} & 0 < a < 1 \\ 1 & a \geq 1 \end{cases}$$

Thus, the density function of Y is

$$f_Y(a) = F'_Y(a) = \begin{cases} \frac{1}{2} a^{-\frac{1}{2}} & 0 < a < 1 \\ 0 & \text{otherwise} \end{cases}$$

Hence, Y is a beta random variable with parameters $(\frac{1}{2}, 1)$
38.16 (a)

$$f_Y(a) = \begin{cases} \frac{a^2}{18} & -3 < a < 3 \\ 0 & \text{otherwise} \end{cases}$$

(b)

$$f_Z(a) = \begin{cases} \frac{3}{2}(3-a)^2 & 2 < a < 4 \\ 0 & \text{otherwise} \end{cases}$$

38.17 $f_Y(y) = \frac{1}{\sqrt{y}} - 1, \ 0 < y \le 1$
38.18 $f_Y(y) = e^{2y - \frac{1}{2}e^{2y}}$
38.19 (a) $f_Y(y) = \frac{8-y}{50}$ for $-2 \le y \le 8$ and 0 elsewhere.
(b) We have $E(Y) = \frac{4}{3}$ (c) $\frac{9}{25}$

Section 39

39.1 (a) From the table we see that the sum of all the entries is 1.
(b) 0.25 (c) 0.55 (d) $p_X(0) = 0.3, p_X(1) = 0.5, p_X(2) = 0.125, p_X(3) = 0.075$ and 0 otherwise
39.2 0.25
39.3 (a) 0.08 (b) 0.36 (c) 0.86 (d) 0.35 (e) 0.6 (f) 0.65 (g) 0.4
39.4 (a) 0.4 (b) 0.8
39.5 0.0423
39.6 (a) The cdf od X and Y is

$$\begin{aligned} F_{XY}(x,y) =& \int_{-\infty}^x \int_{-\infty}^y f_{XY}(u,v)dudv \\ =& \left(\int_0^x ue^{-\frac{u^2}{2}}du\right)\left(\int_0^y ve^{-\frac{v^2}{2}}du\right) \\ =&(1-e^{-\frac{x^2}{2}})(1-e^{-\frac{y^2}{2}}), \ x > 0, y > 0 \end{aligned}$$

and 0 otherwise.
(b) The marginal pdf for X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_0^{\infty} xye^{-\frac{x^2+y^2}{2}}dy = xe^{-\frac{x^2}{2}}, \ x > 0$$

and 0 otherwise. The marginal pdf for Y is

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx = \int_0^{\infty} xye^{-\frac{x^2+y^2}{2}}dx = ye^{-\frac{y^2}{2}}$$

for $y > 0$ and 0 otherwise
39.7 We have

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^a y^{1-a}dxdy = \int_0^1\int_0^1 x^a y^{1-a}dxdy$$
$$=(2+a-a^2)^{-1} \neq 1$$

so $f_{XY}(x,y)$ is not a density function. However one can easily turn it into a density function by multiplying $f(x,y)$ by $(2+a-a^2)$ to obtain the density function

$$f_{XY}(x,y) = \begin{cases} (2+a-a^2)x^a y^{1-a} & 0 \le x, y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

39.8 0.625
39.9 0.708
39.10 0.576
39.11 0.488
39.12 $f_Y(y) = \int_y^{\sqrt{y}} 15y dx = 15y^{\frac{3}{2}}(1 - y^{\frac{1}{2}}), \quad 0 < y < 1$ and 0 otherwise
39.13 5.778
39.14 0.83
39.15 0.008
39.16 $\frac{7}{20}$
39.17 $1 - 2e^{-1}$
39.18 $\frac{12}{25}$
39.19 $\frac{3}{8}$
39.20 (a) We have

X\ Y	1	2	$p_X(x)$
1	0.2	0.5	0.7
2	0.2	0.1	0.3
$p_Y(y)$	0.4	0.6	1

(b) We have

$$F_{XY}(x,y) = \begin{cases} 0 & x < 1 \text{ or } y < 1 \\ 0.2 & 1 \le x < 2 \text{ and } 1 \le y < 2 \\ 0.7 & 1 \le x < 2 \text{ and } y \ge 2 \\ 0.4 & x \ge 2 \text{ and } 1 \le y < 2 \\ 1 & x \ge 2 \text{ and } y \ge 2 \end{cases}$$

39.21 $\int_0^{\frac{1}{2}}\int_{\frac{1}{2}}^1 f(s,t)dsdt + \int_0^1\int_0^{\frac{1}{2}} f(s,t)dsdt$
39.22 $\frac{1}{800}\int_0^{60}\int_0^{60-x} e^{\frac{x}{40}+\frac{y}{20}}dydx$

Section 40

40.1 (a) Yes (b) 0.5 (c) $1-e^{-a}$
40.2 (a) The joint density over the region R must integrate to 1, so we have

$$1=\int\int_{(x,y)\in R} cdxdy = cA(R).$$

(b) Note that $A(R)=4$ so that $f_{XY}(x,y)=\frac{1}{4}=\frac{1}{2}\frac{1}{2}$. Hence, by Theorem 40.2, X and Y are independent with each distributed uniformly over $(-1,1)$.
(c) $\Pr(X^2+Y^2\le 1)=\int\int_{x^2+y^2\le 1}\frac{1}{4}dxdy=\frac{\pi}{4}$
40.3 (a) 0.484375 (b) We have

$$f_X(x)=\int_x^1 6(1-y)dy = 6y-3y^2\Big|_x^1 = 3x^2-6x+3,\ 0\le x\le 1$$

and 0 otherwise. Similarly,

$$f_Y(y)=\int_0^y 6(1-y)dy = 6\left[y-\frac{y^2}{2}\right]_0^y = 6y(1-y),\ 0\le y\le 1$$

and 0 otherwise.
(c) X and Y are dependent
40.4 (a) $k=4$ (b) We have

$$f_X(x)=\int_0^1 4xydy = 2x,\quad 0\le x\le 1$$

and 0 otherwise. Similarly,

$$f_Y(y)=\int_0^1 4xydx = 2y,\quad 0\le y\le 1$$

and 0 otherwise.
(c) Since $f_{XY}(x,y)=f_X(x)f_Y(y)$, X and Y are independent.
40.5 (a) $k=6$ (b) We have

$$f_X(x)=\int_0^1 6xy^2dy = 2x,\quad 0\le x\le 1,\quad 0\ \text{ otherwise}$$

and

$$f_Y(y) = \int_0^1 6xy^2 dy = 3y^2, \quad 0 \le y \le 1, \quad 0 \text{ otherwise}$$

(c) 0.15 (d) 0.875 (e) X and Y are independent
40.6 (a) $k = \frac{8}{7}$ (b) Yes (c) $\frac{16}{21}$
40.7 (a) We have

$$f_X(x) = \int_0^2 \frac{3x^2+2y}{24} dy = \frac{6x^2+4}{24}, 0 \le x \le 2, 0 \text{ otherwise}$$

and

$$f_Y(y) = \int_0^2 \frac{3x^2+2y}{24} dx = \frac{8+4y}{24}, 0 \le y \le 2, \quad 0 \text{ otherwise}$$

(b) X and Y are dependent. (c) 0.340
40.8 (a) We have

$$f_X(x) = \int_x^{3-x} \frac{4}{9} dy = \frac{4}{3} - \frac{8}{9}x, \quad 0 \le x \le \frac{3}{2}, \quad 0 \text{ otherwise}$$

and

$$f_Y(y) = \begin{cases} \frac{4}{9}y & 0 \le y \le \frac{3}{2} \\ \frac{4}{9}(3-y) & \frac{3}{2} \le y \le 3 \\ 0 & \text{otherwise} \end{cases}$$

(b) $\frac{2}{3}$ (c) X and Y are dependent
40.9 0.469
40.10 0.191
40.11 0.4
40.12 0.19
40.13 0.295
40.14 0.414
40.15 $f(z) = e^{-\frac{1}{2}z} - e^{-z}, \quad z > 0, \ 0$ otherwise
40.16 $f(x) = \frac{2}{(2x+1)^2}, \ x > 0, \quad 0$ otherwise
40.17 $\frac{3}{5}$
40.18 Suppose that X and Y are independent. Then $\Pr(X = 0|Y = 1) = \Pr(X = 0) = 0.6$ and $\Pr(X = 1|Y = 0) = 0.7$. Since $\Pr(X = 0) + \Pr(X = 1) = 0.6 + 0.7 \ne 1$, it follows that X and Y can not be independent.

40.19 $\theta_1 = \frac{1}{4}$ and $\theta_2 = 0$

Section 41

41.1

$$\begin{aligned}
\Pr(Z=0)=&\Pr(X=0)\Pr(Y=0)=(0.1)(0.25)=0.025\\
\Pr(Z=1)=&\Pr(X=1)\Pr(Y=0)+\Pr(Y=1)\Pr(X=0)\\
=&(0.2)(0.25)+(0.4)(0.1)=0.09\\
\Pr(Z=2)=&\Pr(X=1)\Pr(Y=1)+\Pr(X=2)\Pr(Y=0)+\Pr(Y=2)\Pr(X=0)\\
=&(0.2)(0.4)+(0.3)(0.25)+(0.35)(0.1)=0.19\\
\Pr(Z=3)=&\Pr(X=2)\Pr(Y=1)+\Pr(Y=2)\Pr(X=1)+\Pr(X=3)\Pr(Y=0)\\
=&(0.3)(0.4)+(0.35)(0.2)+(0.4)(0.25)=0.29\\
\Pr(Z=4)=&\Pr(X=2)\Pr(Y=2)+\Pr(X=3)\Pr(Y=1)\\
=&(0.3)(0.35)+(0.4)(0.4)=0.265\\
\Pr(Z=5)=&\Pr(X=3)\Pr(X=2)=(0.4)(0.35)=0.14
\end{aligned}$$

and 0 otherwise

41.2 $p_{X+Y}(k)=\left(\begin{array}{c}30\\k\end{array}\right)0.2^k0.8^{30-k}$ for $0\le k\le 30$ and 0 otherwise.

41.3 $p_{X+Y}(n)=(n-1)p^2(1-p)^{n-2},\;\; n=2,\cdots$ and $p_{X+Y}(a)=0$ otherwise.

41.4

$$\begin{aligned}
p_{X+Y}(3)=&p_X(0)p_Y(3)=\frac{1}{3}\cdot\frac{1}{4}=\frac{1}{12}\\
p_{X+Y}(4)=&p_X(0)p_Y(4)+p_X(1)p_Y(3)=\frac{4}{12}\\
p_{X+Y}(5)=&p_X(1)p_Y(4)+p_X(2)p_Y(3)=\frac{4}{12}\\
p_{X+Y}(6)=&p_X(2)p_Y(4)=\frac{3}{12}
\end{aligned}$$

and 0 otherwise.

41.5 $\frac{1}{64}$

41.6 0.03368

41.7 $\Pr(X+Y=2)=e^{-\lambda}p(1-p)+e^{-\lambda}\lambda p$ (b) $\Pr(Y>X)=e^{-\lambda p}$

41.8

$$\begin{aligned}
p_{X+Y}(0) =& p_X(0)p_Y(0) = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}\\
p_{X+Y}(1) =& p_X(0)p_Y(1) + p_X(1)p_Y(0) = \frac{1}{2}\cdot\frac{1}{4} + \frac{1}{4}\cdot\frac{1}{2} = \frac{1}{4}\\
p_{X+Y}(2) =& p_X(0)p_Y(2) + p_X(2)p_Y(0) + p_X(1)p_Y(1) = \frac{5}{16}\\
p_{X+Y}(3) =& p_X(1)p_Y(2) + p_X(2)p_Y(1) = \frac{1}{8}\\
p_{X+Y}(4) =& p_X(2)p_Y(2) = \frac{1}{16}
\end{aligned}$$

and 0 otherwise.
41.9

$$\begin{aligned}
p_{X+Y}(1) =& p_X(0)p_Y(1) + p_X(1)p_Y(0) = \frac{1}{6}\\
p_{X+Y}(2) =& p_X(0)p_Y(2) + p_X(2)p_Y(0) + p_X(1)p_Y(1) = \frac{5}{18}\\
p_{X+Y}(3) =& p_X(0)p_Y(3) + p_X(1)p_Y(2) + p_X(2)p_Y(1) + p_X(3)p_Y(0) = \frac{6}{18}\\
p_{X+Y}(4) =& p_X(0)p_Y(4) + p_X(1)p_Y(3) + p_X(2)p_Y(2) + p_X(3)p_Y(1) + p_X(4)p_Y(0) = \frac{3}{18}\\
p_{X+Y}(5) =& p_X(0)p_Y(5) + p_X(1)p_Y(4) + p_X(2)p_Y(3) + p_X(3)p_Y(2)\\
&+ p_X(4)p_Y(1) + p_X(4)p_Y(1) = \frac{1}{18}
\end{aligned}$$

and 0 otherwise.
41.10 We have

$$p_{X+Y}(a) = \sum_{n=0}^{a} p(1-p)^n p(1-p)^{a-n} = (a+1)p^2(1-p)^a = {}_{a+1}C_a p^2(1-p)^a$$

Thus, $X+Y$ is a negative binomial with parameters $(2, p)$.
41.11 $9e^{-8}$
41.12 $e^{-10\lambda}\frac{(10\lambda)^{10}}{10!}$

Section 42

42.1

$$f_{X+Y}(a) = \begin{cases} 2\lambda e^{-\lambda a}(1-e^{-\lambda a}) & 0 \le a\\ 0 & \text{otherwise}\end{cases}$$

42.2

$$f_{X+Y}(a) = \begin{cases} 1-e^{-\lambda a} & 0 \le a \le 1 \\ e^{-\lambda a}(e^{\lambda}-1) & a \ge 1 \\ 0 & \text{otherwise} \end{cases}$$

42.3 $f_{X+2Y}(a) = \int_{-\infty}^{\infty} f_X(a-2y)f_Y(y)dy$
42.4 If $0 \le a \le 1$ then $f_{X+Y}(a) = 2a - \frac{3}{2}a^2 + \frac{a^3}{6}$. If $1 \le a \le 2$ then $f_{X+Y}(a) = \frac{7}{6} - \frac{a}{2}$. If $2 \le a \le 3$ then $f_{X+Y}(a) = \frac{9}{2} - \frac{9}{2}a + \frac{3}{2}a^2 - \frac{1}{6}a^3$. If $a > 3$ then $f_{X+Y}(a) = 0$.
42.5 If $0 \le a \le 1$ then $f_{X+Y}(a) = \frac{2}{3}a^3$. If $1 < a < 2$ then $f_{X+Y}(a) = -\frac{2}{3}a^3 + 4a - \frac{8}{3}$. If $a \ge 2$ then $f_{X+Y}(a) = 0$ and 0 otherwise.
42.6 $f_{X+Y}(a) = \frac{\alpha\beta}{\alpha-\beta}\left(e^{-\beta a} - e^{-\alpha a}\right)$ for $a > 0$ and 0 otherwise.
42.7 $f_W(a) = e^{-\frac{a}{2}} - e^{-a}, \quad a > 0$ and 0 otherwise.
42.8 If $2 \le a \le 4$ then $f_{X+Y}(a) = \frac{a}{4} - \frac{1}{2}$. If $4 \le a \le 6$, then $f_{X+Y}(a) = \frac{3}{2} - \frac{a}{4}$ and $f_{X+Y}(a) = 0$ otherwise.
42.9 If $0 < a \le 2$ then $f_{X+Y}(a) = \frac{a^2}{8}$. If $2 < a < 4$ then $f_{X+Y}(a) = -\frac{a^2}{8} + \frac{a}{2}$ and 0 otherwise.
42.10 $f_{X+Y}(a) = \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} e^{-(a-(\mu_1+\mu_2))^2/[2(\sigma_1^2+\sigma_2^2)]}$.
42.11 $\frac{1}{8}$
42.12 $f_Z(z) = \int_0^z e^{-z}ds = ze^{-z}$ for $z > 0$ and 0 otherwise.
42.13 $1 - 2e^{-1}$

Section 43

43.1 $p_{X|Y}(0|1) = 0.25$ and $p_{X|Y}(1|1) = 0.75$ and 0 otherwise.
43.2 (a) For $1 \le x \le 5$ and $y = 1, \cdots, x$ we have $p_{XY}(x,y) = \left(\frac{1}{5}\right)\left(\frac{1}{x}\right)$ and 0 otherwise.
(b) $p_{X|Y}(x|y) = \frac{\frac{1}{5x}}{\sum_{k=y}^{5}\left(\frac{1}{5k}\right)}$ and 0 otherwise.
(c) X and Y are dependent

43.3

$$\begin{aligned}
\Pr(X=3|Y=4) =& \frac{\Pr(X=3,Y=4)}{\Pr(Y=4)}=\frac{0.10}{0.35}=\frac{2}{7}\\
\Pr(X=4|Y=4) =& \frac{\Pr(X=4,Y=4)}{\Pr(Y=4)}=\frac{0.15}{0.35}=\frac{3}{7}\\
\Pr(X=5|Y=4) =& \frac{\Pr(X=5,Y=4)}{\Pr(Y=4)}=\frac{0.10}{0.35}=\frac{2}{7}
\end{aligned}$$

43.4

$$\begin{aligned}
\Pr(X=0|Y=1) =& \frac{\Pr(X=0,Y=1)}{\Pr(Y=1)}=\frac{1/16}{6/16}=\frac{1}{6}\\
\Pr(X=1|Y=1) =& \frac{\Pr(X=1,Y=1)}{\Pr(Y=1)}=\frac{3/16}{6/16}=\frac{1}{2}\\
\Pr(X=2|Y=1) =& \frac{\Pr(X=2,Y=1)}{\Pr(Y=1)}=\frac{2/16}{6/16}=\frac{1}{3}\\
\Pr(X=3|Y=1) =& \frac{\Pr(X=3,Y=1)}{\Pr(Y=1)}=\frac{0/16}{6/16}=0
\end{aligned}$$

and 0 otherwise.
43.5

y	1	2	3	4	5	6
$p_{Y\|X}(1\|x))$	1	$\frac{2}{3}$	$\frac{2}{5}$	$\frac{2}{7}$	$\frac{2}{9}$	$\frac{2}{11}$
$p_{Y\|X}(2\|x)$	0	$\frac{1}{3}$	$\frac{2}{5}$	$\frac{2}{7}$	$\frac{2}{9}$	$\frac{2}{11}$
$p_{Y\|X}(3\|x)$	0	0	$\frac{1}{5}$	$\frac{2}{7}$	$\frac{2}{9}$	$\frac{2}{11}$
$p_{Y\|X}(4\|x)$	0	0	0	$\frac{1}{7}$	$\frac{2}{9}$	$\frac{2}{11}$
$p_{Y\|X}(5\|x)$	0	0	0	0	$\frac{1}{9}$	$\frac{2}{11}$
$p_{Y\|X}(6\|x)$	0	0	0	0	0	$\frac{1}{11}$

and 0 otherwise. X and Y are dependent since $p_{Y|X}(1|1)=1\neq\frac{11}{36}=p_Y(1)$.
43.6 (a) $p_Y(y)={}_nC_yp^y(1-p)^{n-y}$ and 0 otherwise. Thus, Y is a binomial distribution with parameters n and p.

(b)

$$\begin{aligned}
p_{X|Y}(x|y) =& \frac{p_{XY}(x,y)}{p_Y(y)} \\
=& \frac{\frac{n!y^x(pe^{-1})^y(1-p)^{n-y}}{y!(n-y)!x!}}{{}_nC_y p^y(1-p)^{n-y}} \\
=& \frac{y^x e^{-y}}{x!}, \quad x = 0,1,2,\cdots \\
=& 0, \quad \text{otherwise}
\end{aligned}$$

Thus, $X|Y=y$ is a Poisson distribution with parameter y.
X and Y are not independent.
43.7

$$p_{X|Y}(x|0) = \frac{p_{XY}(x,0)}{p_Y(0)} = \begin{cases} 1/11 & x=0 \\ 4/11 & x=1 \\ 6/11 & x=2 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{X|Y}(x|1) = \frac{p_{XY}(x,1)}{p_Y(1)} = \begin{cases} 3/7 & x=0 \\ 3/7 & x=1 \\ 1/7 & x=2 \\ 0 & \text{otherwise} \end{cases}$$

The conditional probability distribution for Y given $X=x$ is

$$p_{Y|X}(y|0) = \frac{p_{XY}(0,y)}{p_X(0)} = \begin{cases} 1/4 & y=0 \\ 3/4 & y=1 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{Y|X}(y|1) = \frac{p_{XY}(1,y)}{p_X(1)} = \begin{cases} 4/7 & y=0 \\ 3/7 & y=1 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{Y|X}(y|2) = \frac{p_{XY}(2,y)}{p_X(2)} = \begin{cases} 6/7 & y=0 \\ 1/7 & y=1 \\ 0 & \text{otherwise} \end{cases}$$

43.8 (a) $\frac{1}{2^N-1}$ (b) $p_X(x) = \frac{2^x}{2^N-1}$ for $x = 0,1,\cdots,N-1$ and 0 otherwise. (c) $p_{Y|X}(y|x) = 2^{-x}(1-2^{-x})^y$ for $x = 0,1,\cdots,N-1$, $y = 0,1,2,\cdots$ and 0 otherwise

43.9 $\Pr(X = k|X + Y = n) = C(n,k)\left(\frac{1}{2}\right)^n$ for $k = 0, 1, \cdots, n$ and 0 otherwise.
43.10 (a)

$$\begin{aligned}
\Pr(X = 0, Y = 0) =& \frac{48}{52}\frac{47}{51} = \frac{188}{221}\\
\Pr(X = 1, Y = 0) =& \frac{48}{52}\frac{4}{51} = \frac{16}{221}\\
\Pr(X = 1, Y = 1) =& \frac{4}{52}\frac{48}{51} = \frac{16}{221}\\
\Pr(X = 2, Y = 1) =& \frac{4}{52}\frac{3}{51} = \frac{1}{221}
\end{aligned}$$

and 0 otherwise.
(b) $\Pr(Y = 0) = \Pr(X = 0, Y = 0) + \Pr(X = 1, Y = 0) = \frac{204}{221} = \frac{12}{13}$ and $\Pr(Y = 1) = \Pr(X = 1, Y = 1) + \Pr(X = 2, Y = 2) = \frac{1}{13}$ and 0 otherwise.
(c) $p_{X|Y}(1|1) = 13 \times \frac{16}{221} = \frac{16}{17}$ and $p_{X|Y}(2|1) = 13 \times \frac{1}{221} = \frac{1}{17}$ and 0 otherwise
43.11 e^2

Section 44

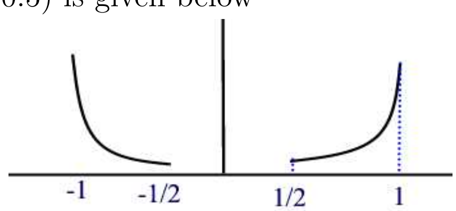
44.1 For $y \le |x| \le 1, 0 \le y \le 1$ we have

$$f_{X|Y}(x|y) = \frac{3}{2}\left[\frac{x^2}{1 - y^3}\right]$$

If $y = 0.5$ then

$$f_{X|Y}(x|0.5) = \frac{12}{7}x^2, \quad 0.5 \le |x| \le 1.$$

The graph of $f_{X|Y}(x|0.5)$ is given below



44.2 $f_{X|Y}(x|y) = \frac{2x}{y^2}, \quad 0 \le x < y \le 1$
44.3 $f_{Y|X}(y|x) = \frac{3y^2}{x^3}, \quad 0 \le y < x \le 1$
44.4 $f_{X|Y}(x|y) = (y + 1)^2 x e^{-x(y+1)}, \ x \ge 0$ and $f_{Y|X}(x|y) = xe^{-xy}, \ y \ge 0$

44.5 (a) For $0 < y < x$ we have $f_{XY}(x,y) = \frac{y^2}{2}e^{-x}$
(b) $f_{X|Y}(x|y) = e^{-(x-y)}, \quad 0 < y < x$
44.6 (a) $f_{X|Y}(x|y) = 6x(1-x), \quad 0 < x < 1$. X and Y are independent.
(b) 0.25
44.7 $f_{X|Y}(x|y) = \frac{\frac{1}{3}x-y+1}{\frac{3}{2}-y}$ (b) $\frac{11}{24}$
44.8 0.25
44.9 $\frac{8}{9}$
44.10 0.4167
44.11 $\frac{7}{8}$
44.12 0.1222
44.13 $f_X(x) = \int_{x^2}^1 \frac{1}{\sqrt{y}}dy = 2(1-x), \quad 0 < x < 1$ and 0 otherwise
44.14 $\frac{1}{1-y}$ for $0 < y < x < 1$
44.15 mean$=\frac{1}{3}$ and $\mathrm{Var}(Y) = \frac{1}{18}$
44.16 0.172

Section 45

45.1 $f_{ZW}(z,w) = \frac{f_{XY}(\frac{zd-bw}{ad-bc},\frac{aw-cz}{ad-bc})}{|ad-bc|}$
45.2 $f_{Y_1Y_2}(y_1,y_2) = \frac{\lambda^2}{y_2}e^{-\lambda y_1}, \ y_2 > 1, \ y_1 \geq \ln y_2$
45.3 $f_{R\Phi}(r,\phi) = rf_{XY}(r\cos\phi, r\sin\phi), \quad r > 0, \quad -\pi < \phi \leq \pi$
45.4

$$\begin{aligned} f_{ZW}(z,w) =& \frac{z}{1+w^2}[f_{XY}(z(\sqrt{1+w^2})^{-1}, wz(\sqrt{1+w^2})^{-1}) \\ &+ f_{XY}(-z(\sqrt{1+w^2})^{-1}, -wz(\sqrt{1+w^2})^{-1})] \end{aligned}$$

45.5 $f_{UV}(u,v) = \frac{\lambda e^{-\lambda u}(\lambda u)^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)}\frac{v^{\alpha-1}(1-v)^{\beta-1}\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$. Hence $X+Y$ and $\frac{X}{X+Y}$ are independent, with $X+Y$ having a gamma distribution with parameters $(\alpha+\beta,\lambda)$ and $\frac{X}{X+Y}$ having a beta distribution with parameters (α,β).
45.6

$$f_{Y_1Y_2}(y_1,y_2) = \begin{cases} e^{-y_1}y_1 & y_1 \geq 0, \ 0 < y_2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

45.7 $f_{Y_1Y_2}(y_1,y_2) = \frac{1}{\sqrt{2\pi}}e^{-(\frac{3}{7}y_1+\frac{1}{7}y_2)^2/2}\frac{1}{\sqrt{8\pi}}e^{-(\frac{1}{7}y_1-\frac{2}{7}y_2)^2/8} \cdot \frac{1}{7}$
45.8 We have $u = g_1(x,y) = \sqrt{2y}\cos x$ and $v = g_2(x,y) = \sqrt{2y}\sin x$. The Jacobian of the transformation is

$$J = \begin{vmatrix} -\sqrt{2y}\sin x & \frac{\cos x}{\sqrt{2y}} \\ \sqrt{2y}\cos x & \frac{\sin x}{\sqrt{2y}} \end{vmatrix} = -1$$

Also,

$$y = \frac{u^2 + v^2}{2}$$

Thus,

$$f_{UV}(u,v) = f_{XY}(x(u,v), y(u,v))|J|^{-1} = \frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}.$$

This is the joint density of two independent standard normal random variables.

45.9 $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_{XY}(a-y,y)dy$ If X and Y are independent then $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy$

45.10 $f_{Y-X}(a) = \int_{-\infty}^{\infty} f_{XY}(y-a,y)dy$. If X and Y are independent then $f_{Y-X}(a) = \int_{-\infty}^{\infty} f_X(y-a)f_Y(y)dy = \int_{-\infty}^{\infty} f_X(y)f_Y(a+y)dy$

45.11 $f_U(u) = \int_{-\infty}^{\infty} \frac{1}{|v|}f_{XY}(v, \frac{u}{v})dv$. If X and Y are independent then $f_U(u) = \int_{-\infty}^{\infty} \frac{1}{|v|}f_X(v)f_Y(\frac{u}{v})dv$

45.12 $f_U(u) = \frac{1}{(u+1)^2}$ for $u > 0$ and 0 elsewhere.

Section 46

46.1 $\frac{(m+1)(m-1)}{3m}$ **46.2** $E(XY) = \frac{7}{12}$

46.3 $E(|X-Y|) = \frac{1}{3}$

46.4 $E(X^2Y) = \frac{7}{36}$ and $E(X^2+Y^2) = \frac{5}{6}$.

46.5 0

46.6 33

46.7 $\frac{L}{3}$

46.8 $\frac{30}{19}$

46.9 (a) 0.9 (b) 4.9 (c) 4.2

46.10 (a) 14 (b) 45

46.11 5.725

46.12 $\frac{2}{3}L^2$

46.13 27

46.14 5

Section 47

47.1 $2\sigma^2$

47.2 coveraince is -0.123 and correlation is -0.33

47.3 covariance is $\frac{1}{12}$ and correlation is $\frac{\sqrt{2}}{2}$

47.4 (a) $f_{XY}(x,y) = 5, \quad -1 < x < 1,\ x^2 < y < x^2 + 0.1$ and 0 otherwise.
(b) covariance is 0 and correlation is 0
47.5 We have

$$E(X) = \frac{1}{2\pi}\int_0^{2\pi} \cos\theta d\theta = 0$$

$$E(Y) = \frac{1}{2\pi}\int_0^{2\pi} \sin\theta d\theta = 0$$

$$E(XY) = \frac{1}{2\pi}\int_0^{2\pi} \cos\theta \sin\theta d\theta = 0$$

Thus X and Y are uncorrelated, but they are clearly not independent, since they are both functions of θ
47.6 (a) $\rho(X_1+X_2, X_2+X_3) = 0.5$ (b) $\rho(X_1+X_2, X_3+X_4) = 0$
47.7 $-\frac{n}{36}$
47.8 We have

$$E(X) = \frac{1}{2}\int_{-1}^{1} x dx = 0$$

$$E(XY) = E(X^3) = \frac{1}{2}\int_{-1}^{1} x^3 dx = 0$$

Thus, $\rho(X,Y) = Cov(X,Y) = 0$
47.9 $Cov(X,Y) = \frac{3}{160}$ and $\rho(X,Y) = 0.397$
47.10 24
47.11 19,300
47.12 11
47.13 200
47.14 0
47.15 0.04
47.16 6
47.17 8.8
47.18 0.2743
47.19 (a)

$$f_{XY}(x,y) = f_X(x)f_Y(y) = \begin{cases} \frac{1}{2} & 0 < x < 1, 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

(b)

$$f_Z(a) = \begin{cases} 0 & a \le 0 \\ \frac{a}{2} & 0 < a \le 1 \\ \frac{1}{2} & 1 < a \le 2 \\ \frac{3-a}{2} & 2 < a \le 3 \\ 0 & a > 3 \end{cases}$$

(c) $E(X) = 0.5, \ \mathrm{Var}(X) = \frac{1}{12}, \ E(Y) = 1, \ \mathrm{Var}(Y) = \frac{1}{3}$
(d) $E(Z) = 1.5$ and $\mathrm{Var}(Z) = \frac{5}{12}$
47.20 (a) $f_Z(z) = \frac{dF_z}{dz}(z) = \frac{z}{2}, \ 0 < z < 2$ and 0 otherwise.
(b) $f_X(x) = 1 - \frac{x}{2}, \quad 0 < x < 2$ and 0 otherwise; $f_Y(y) = 1 - \frac{y}{2}, \quad 0 < y < 2$ and 0 otherwise.
(c) $E(X) = \frac{2}{3}$ and $\mathrm{Var}(X) = \frac{2}{9}$
(d) $\mathrm{Cov}(X, Y) = -\frac{1}{9}$
47.21 -0.15
47.22 $\frac{1}{6}$
47.23 $\frac{5}{12}$
47.24 1.04
47.25 $\frac{\sqrt{\pi}}{2}$
47.26 $E(W) = 4$ and $\mathrm{Var}(W) = 67$
47.27 $-\frac{1}{5}$
47.28 2
47.29 $\frac{n-2}{n+2}$
47.30 (a) We have

X \ Y	0	1	2	$p_X(x)$
0	0.25	0.08	0.05	0.38
1	0.12	0.20	0.10	0.42
2	0.03	0.07	0.10	0.2
$P_Y(y)$	0.4	0.35	0.25	1

(b) $E(X) = 0.82$ and $E(Y) = 0.85$.
(c) 0.243 (d) 145.75

Section 48

48.1 $E(X|Y) = \frac{2}{3}Y$ and $E(Y|X) = \frac{2}{3}\left(\frac{1-X^3}{1-X^2}\right)$
48.2 $E(X) = \frac{1}{2}, \ E(X^2) = \frac{1}{3}, \ Var(X) = \frac{1}{12}, \ E(Y|X) = \frac{3}{4}X, \ Var(Y|X) = \frac{3}{80}X^2, \ E[Var(Y|X)] = \frac{1}{80}, \ Var[E(Y|X)] = \frac{3}{64}, \ \mathrm{Var}(Y) = \frac{19}{320}$

48.3 $\frac{\lambda}{\lambda+\mu}$ **48.4** 0.75
48.5 2.3
48.6 $\frac{2}{3}\left(\frac{1-X^6}{1-X^4}\right)$
48.7 A first way for finding $E(Y)$ is

$$E(Y)=\int_0^1 y\frac{7}{2}y^{\frac{5}{2}}dy=\int_0^1 y\frac{7}{2}y^{\frac{7}{2}}dy=\frac{7}{9}.$$

For the second way, we use the double expectation result

$$E(Y)=E(E(Y|X))=\int_{-1}^1 E(Y|X)f_X(x)dx=\int_{-1}^1\frac{2}{3}\left(\frac{1-x^6}{1-x^4}\right)\frac{21}{8}x^2(1-x^6)=\frac{7}{9}$$

48.8 $E(X)=15$ and $E(Y)=5$
48.9 -1.4 **48.10** 15
48.11 $E(X|Y=1)=2,\ E(X|Y=2)=\frac{5}{3},\ E(X|Y=3)=\sum_x\frac{p_{XY}(x,1)}{p_Y(3)}=\frac{12}{5}.$
X and Y are dependent.
48.12 0.20
48.13 $\frac{1}{12}$
48.14 0.9856
48.15 13
48.16 $\frac{(1-x)^2}{12}$
48.17 Mean is $\alpha\lambda$ and variance is $\beta^2(\lambda+\lambda^2)+\alpha^2\lambda$
48.18 0.25
48.19 2.25
48.20 6.6
48.21 0.534
48.22 0.076
48.23 0.0756
48.24 8,000,000
48.25 $\frac{2}{3}\frac{1+x+x^2}{1+x}$

Section 49

49.1 $E(X)=\frac{n+1}{2}$ and $\text{Var}(X)=\frac{n^2-1}{12}$
49.2 $E(X)=\frac{1}{p}$ and $\text{Var}(X)=\frac{1-p}{p^2}$
49.3 The moment generating function is

$$M_X(t)=\sum_{n=1}^\infty e^{tn}\frac{6}{\pi^2n^2}.$$

By the ratio test we have

$$\lim_{n\to\infty}\frac{e^{t(n+1)}\frac{6}{\pi^2(n+1)^2}}{e^{tn}\frac{6}{\pi^2n^2}} = \lim_{n\to\infty} e^t\frac{n^2}{(n+1)^2} = e^t > 1$$

and so the summation diverges whenever $t > 0$. Hence there does not exist a neighborhood about 0 in which the mgf is finite.
49.4 $E(X) = \frac{\alpha}{\lambda}$ and $\text{Var}(X) = \frac{\alpha}{\lambda^2}$
49.5 Let $Y = X_1 + X_2 + \cdots + X_n$ where each X_i is an exponential random variable with parameter λ. Then

$$M_Y(t) = \prod_{k=1}^{n} M_{X_k}(t) = \prod_{k=1}^{n}\left(\frac{\lambda}{\lambda - t}\right) = \left(\frac{\lambda}{\lambda - t}\right)^n, \quad t < \lambda.$$

Since this is the mgf of a gamma random variable with parameters n and λ we can conclude that Y is a gamma random variable with parameters n and λ.
49.6 $M_X(t) = \begin{cases} 1 & t = 0 \\ \infty & \text{otherwise} \end{cases}$ $\quad \{t \in \mathbb{R} : M_X(t) < \infty\} = \{0\}$
49.7 $M_Y(t) = E(e^{tY}) = e^{-2t}\frac{\lambda}{\lambda - 3t}, \quad 3t < \lambda$
49.8 This is a binomial random variable with $p = \frac{3}{4}$ and $n = 15$
49.9 Y has the same distribution as $3X - 2$ where X is a binomial distribution with $n = 15$ and $p = \frac{3}{4}$.
49.10 $E(t_1W + t_2Z) = e^{\frac{(t_1+t_2)^2}{2}}e^{\frac{(t_1-t_2)^2}{2}} = e^{t_1^2+t_2^2}$
49.11 5,000
49.12 10,560
49.13 $M(t) = E(e^{ty}) = \frac{19}{27} + \frac{8}{27}e^t$
49.14 0.84
49.15 (a) $M_{X_i}(t) = \frac{pe^t}{1-(1-p)e^t}, \quad t < -\ln(1-p)$
(b) $M_X(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^n, \quad t < -\ln(1-p)$.
(c) Because $X_1, X_2, \cdots, X_n$ are independent then

$$\begin{aligned} M_Y(t) = \prod_{k=1}^{n} M_{X_i}(t) = &\prod_{k=1}^{n}\frac{pe^t}{1-(1-p)e^t} \\ &= \left(\frac{pe^t}{1-(1-p)e^t}\right)^n \end{aligned}$$

Because $\left(\frac{pe^t}{1-(1-p)e^t}\right)^n$ is the moment generating function of a negative binomial random variable with parameters (n, p) then $X_1 + X_2 + \cdots + X_n$ is a negative binomial random variable with the same pmf
49.16 0.6915
49.17 2
49.18 $M_X(t) = \frac{e^t(6-6t+3t^2)-6}{t^3}$
49.19 -38
49.20 $\frac{1}{2e}$
49.21 $e^{\frac{k^2}{2}}$
49.22 4
49.23 $\frac{(e^{t_1}-1)(e^{t_2}-1)}{t_1t_2}$
49.24 $\frac{2}{9}$
49.25 e^{13t^2+4t}
49.26 0.4
49.27 $-\frac{15}{16}$
49.28 $(0.7+0.3e^t)^9$
49.29 41.9
49.30 0.70

Section 50

50.1 Clearly $E(X) = -\frac{\epsilon}{2} + \frac{\epsilon}{2} = 0, E(X^2) = \epsilon^2$ and $Var(X) = \epsilon^2$. Thus, $\Pr(|X-0| \geq \epsilon) = 1 = \frac{\sigma^2}{\epsilon^2} = 1$
50.2 100
50.3 0.4444
50.4 $\Pr(X \geq 10^4) \leq \frac{10^3}{10^4} = 0.1$
50.5 $\Pr(0 < X < 40) = \Pr(|X-20| < 20) = 1 - \Pr(|X-20| \geq 20) \geq 1 - \frac{20}{20^2} = \frac{19}{20}$
50.6 $\Pr(X_1 + X_2 + \cdots + X_{20} > 15) \leq 1$
50.7 $\Pr(|X-75| \leq 10) \geq 1 - \Pr(|X-75| \geq 10) \geq 1 - \frac{25}{100} = \frac{3}{4}$
50.8 Using Markov's inequality we find

$$\Pr(X \geq \epsilon) = \Pr(e^{tX} \geq e^{\epsilon t}) \leq \frac{E(e^{tX})}{e^{\epsilon t}},\ t > 0$$

50.9 $\Pr(X > 75) = \Pr(X \geq 76) \leq \frac{50}{76} \approx 0.658$
50.10 $\Pr(0.475 \leq X \leq 0.525) = \Pr(|X-0.5| \leq 0.025) \geq 1 - \frac{25\times 10^{-7}}{625\times 10^{-6}} = 0.996$

50.11 By Markov's inequality $\Pr(X \geq 2\mu) \leq \frac{E(X)}{2\mu} = \frac{\mu}{2\mu} = \frac{1}{2}$
50.12 $\frac{100}{121}$; $\Pr(|X-100| \geq 30) \leq \frac{\sigma^2}{30^2} = \frac{1}{180}$ and $\Pr(|X-100| < 30) \geq 1 - \frac{1}{180} = \frac{179}{180}$. Therefore, the probability that the factory's production will be between 70 and 130 in a day is not smaller than $\frac{179}{180}$
50.13 $\Pr(100 \leq X \leq 140) = 1 - \Pr(|X-120| \geq 21) \geq 1 - \frac{84}{21^2} \approx 0.810$

50.14 We have $E(X) = \int_0^1 x(2x)dx = \frac{2}{3} < \infty$ and $E(X^2) = \int_0^1 x^2(2x)dx = \frac{1}{2} < \infty$ so that $\text{Var}(X) = \frac{1}{2} - \frac{4}{9} = \frac{1}{18} < \infty$. Thus, by the Weak Law of Large Numbers we know that $\overline{X}$ converges in probability to $E(X) = \frac{2}{3}$
50.15 (a)

$$\begin{aligned}
F_{Y_n}(x) =& \Pr(Y_n \leq x) = 1 - \Pr(Y_n > x)\\
=& 1 - \Pr(X_1 > x, X_2 > x, \cdots, X_n > x)\\
=& 1 - \Pr(X_1 > x)\Pr(X_2 > x)\cdots\Pr(X_n > x)\\
=& 1 - (1-x)^n
\end{aligned}$$

for $0 < x < 1$. Also, $F_{Y_n}(x) = 0$ for $x \leq 0$ and $F_{Y_n}(x) = 1$ for $x \geq 1$.
(b) Let $\epsilon > 0$ be given. Then

$$\Pr(|Y_n - 0| \leq \epsilon) = \Pr(Y_n \leq \epsilon) = \begin{cases} 1 & \epsilon \geq 1 \\ 1-(1-\epsilon)^n & 0 < \epsilon < 1 \end{cases}$$

Considering the non-trivial case $0 < \epsilon < 1$ we find

$$\lim_{n\to\infty} \Pr(|Y_n - 0| \leq \epsilon) = \lim_{n\to\infty}[1-(1-\epsilon)^n] = 1 - \lim_{n\to\infty} 1 - 0 = 1.$$

Hence, $Y_n \to 0$ in probability.

Section 51

51.1 0.2119
51.2 0.9876
51.3 0.0094
51.4 0.692
51.5 0.1367
51.6 0.383
51.7 0.0088 **51.8** 0
51.9 23

51.10 6,342,637.5
51.11 0.8185
51.12 16
51.13 0.8413
51.14 0.1587
51.15 0.9887
51.16 (a) $\overline{X}$ is approximated by a normal distribution with mean 100 and variance $\frac{400}{100} = 4$. (b) 0.9544.
51.17 (a) 0.79 (b) 0.9709

Section 52

52.1 (a) 0.167 (b) 0.5833 (c) 0.467 (d) 0.318
52.2 For $t > 0$ we have $\Pr(X \geq a) \leq e^{-ta}(pe^t + 1 - p)^n$ and for $t < 0$ we have $\Pr(X \leq a) \leq e^{-ta}(pe^t + 1 - p)^n$
52.3 0.692
52.4 0.0625
52.5 For $t > 0$ we have $\Pr(X \geq n) \leq e^{-nt}e^{\lambda(e^t-1)}$ and for $t < 0$ we have $\Pr(X \leq n) \leq e^{-nt}e^{\lambda(e^t-1)}$
52.6 (a) 0.769 (b) $\Pr(X \geq 26) \leq e^{-26t}e^{20(e^t-1)}$ (c) 0.357 (d) 0.1093
52.7 Follow from Jensen's inequality
52.8 (a) $\frac{a}{a-1}$ (b) $\frac{a}{a+1}$
(c) We have $g'(x) = -\frac{1}{x^2}$ and $g''(x) = \frac{2}{x^3}$. Since $g''(x) > 0$ for all x in $(0, \infty)$ we conclude that $g(x)$ is convex there.
(d) We have

$$\frac{1}{E(X)} = \frac{a-1}{a} = \frac{a^2-1}{a(a+1)}$$

and

$$E(\frac{1}{X}) = \frac{a}{a+1} = \frac{a^2}{a(a+1)}.$$

Since $a^2 \geq a^2 - 1$, we have $\frac{a^2}{a(a+1)} \geq \frac{a^2-1}{a(a+1)}$. That is, $E(\frac{1}{X}) \geq \frac{1}{E(X)})$, which verifies Jensen's inequality in this case.
52.9 Let X be a random variable such that $\Pr(X = x_i) = \frac{1}{n}$ for $1 \leq i \leq n$. Let $g(x) = \ln x^2$. By Jensen's inequality we have for $X > 0$

$$E[-\ln(X^2)] \geq -\ln[E(X^2)].$$

That is

$$E[\ln(X^2)] \le \ln[E(X^2)].$$

But

$$E[\ln(X^2)] = \frac{1}{n}\sum_{i=1}^{n}\ln x_i^2 = \frac{1}{n}\ln(x_1 \cdot x_2 \cdot \cdots \cdot x_n)^2.$$

and

$$\ln[E(X^2)] = \ln\left(\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}\right)$$

It follows that

$$\ln(x_1 \cdot x_2 \cdot \cdots \cdot x_n)^{\frac{2}{n}} \le \ln\left(\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}\right)$$

or

$$(x_1 \cdot x_2 \cdot \cdots \cdot x_n)^{\frac{2}{n}} \le \frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}$$

Section 53

53.1 (a) 3420 (b) 4995 (c) 5000
53.2 (a)

$$f(y) = \begin{cases} 0.80 & y = 0, x = 0 \\ 20\%(0.50) = 0.10 & y = 0, x = 500 \\ 20\%(0.40) = 0.08 & y = 1800, x = 5000 \\ 20\%(0.10) = 0.02 & y = 5800, x = 15000 \end{cases}$$

(b) 260 (c) 929.73 (d) 490 (e) 1515.55 (f) 0.9940
53.3 0.8201
54.4 1% reduction on the variance

Bibliography

[1] Sheldon Ross, *A first course in Probability*, 8th Edition (2010), Prentice Hall.

[2] SOA/CAS, *Exam P Sample Questions*

Index