

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

Marcel B. Finan
Arkansas Tech University
©All Rights Reserved
Preliminary Draft

In memory of my mother

August 1, 2008

Contents

Preface	7
Risk Management Concepts	8
Basic Operations on Sets	11
1 Basic Definitions	12
2 Set Operations	19
Counting and Combinatorics	33
3 The Fundamental Principle of Counting	33
4 Permutations and Combinations	39
5 Permutations and Combinations with Indistinguishable Objects	49
Probability: Definitions and Properties	59
6 Basic Definitions and Axioms of Probability	59
7 Properties of Probability	67
8 Probability and Counting Techniques	75
Conditional Probability and Independence	81
9 Conditional Probabilities	81
10 Posterior Probabilities: Bayes' Formula	89
11 Independent Events	100
12 Odds and Conditional Probability	109
Discrete Random Variables	113
13 Random Variables	113
14 Probability Mass Function and Cumulative Distribution Function	118
15 Expected Value of a Discrete Random Variable	126
16 Expected Value of a Function of a Discrete Random Variable	133

17 Variance and Standard Deviation	140
18 Binomial and Multinomial Random Variables	146
19 Poisson Random Variable	160
20 Other Discrete Random Variables	170
20.1 Geometric Random Variable	170
20.2 Negative Binomial Random Variable	177
20.3 Hypergeometric Random Variable	184
21 Properties of the Cumulative Distribution Function	190
Continuous Random Variables	205
22 Distribution Functions	205
23 Expectation, Variance and Standard Deviation	217
24 The Uniform Distribution Function	235
25 Normal Random Variables	240
26 Exponential Random Variables	255
27 Gamma and Beta Distributions	265
28 The Distribution of a Function of a Random Variable	277
Joint Distributions	285
29 Jointly Distributed Random Variables	285
30 Independent Random Variables	299
31 Sum of Two Independent Random Variables	310
31.1 Discrete Case	310
31.2 Continuous Case	315
32 Conditional Distributions: Discrete Case	324
33 Conditional Distributions: Continuous Case	331
34 Joint Probability Distributions of Functions of Random Variables	340
Properties of Expectation	347
35 Expected Value of a Function of Two Random Variables	347
36 Covariance, Variance of Sums, and Correlations	357
37 Conditional Expectation	370
38 Moment Generating Functions	381
Limit Theorems	397
39 The Law of Large Numbers	397
39.1 The Weak Law of Large Numbers	397
39.2 The Strong Law of Large Numbers	403

<i>CONTENTS</i>	5
40 The Central Limit Theorem	414
41 More Useful Probabilistic Inequalities	424
Appendix	431
42 Improper Integrals	431
43 Double Integrals	438
44 Double Integrals in Polar Coordinates	451
BIBLIOGRAPHY	456

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 samples of the Exam P/1 provided by the Casual Society of Actuaries will be indicated by the symbol ‡.

This manuscript is also suitable for a one semester course in an undergraduate course in probability theory. Answer keys to text problems can be found by accessing: <http://progressofliberty.today.com/finan-1-answers/>.

A complete solution manual is available to instructors using the book. Email: mfinan@atu.edu

This project has been partially supported by a research grant from Arkansas Tech University.

Marcel B. Finan
Russellville, Ar
May 2007

Risk Management Concepts

When someone is subject to the risk of incurring a financial loss, the loss is generally modeled using a random variable or some combination of random variables. The loss is often related to a particular time interval—for example, an individual may own property that might suffer some damage during the following year. Someone who is at risk of a financial loss may choose some form of **insurance** protection to reduce the impact of the loss. An insurance policy is a contract between the party that is at risk (the **policyholder**) and an **insurer**. This contract generally calls for the policyholder to pay the insurer some specified amount, the **insurance premium**, and in return, the insurer will reimburse certain **claims** to the policyholder. A claim is all or part of the loss that occurs, depending on the nature of the insurance contract.

There are a few ways of modeling a random loss/claim for a particular insurance policy, depending on the nature of the loss. Unless indicated otherwise, we will assume the amount paid to the policyholder as a claim is the amount of the loss that occurs. Once the random variable X representing the loss has been determined, the expected value of the loss, $E(X)$ is referred to as the pure premium for the policy. $E(X)$ is also the expected claim on the insurer. Note that in general, X might be 0 - it is possible that no loss occurs. For a random variable X a measure of the risk is the variation of X , $\sigma^2 = \text{Var}(X)$ to be introduced in Section 17. The **unitized risk** or the **coefficient of variation** for the random variable X is defined to be

$$\frac{\sqrt{\text{Var}(X)}}{E(X)}$$

Partial insurance coverage: It is possible to construct an insurance policy in which the claim paid by the insurer is part, but not necessarily all, of the loss that occurs. There are a few standard types of partial insurance coverage on a basic **ground up loss** random variable X .

(i) **Excess-of-loss insurance:** An excess-of-loss insurance specifies a **deductible amount**, say d . If a loss of amount X occurs, the insurer pays nothing if the loss is less than d , and pays the policyholder the amount of the loss in excess of d if the loss is greater than d .

Two variations on the notion of deductible are

(a) the **franchise deductible**: a franchise deductible of amount d refers to the situation in which the insurer pays 0 if the loss is below d but pays the full amount of loss if the loss is above d ;

(b) the **disappearing deductible**: a disappearing deductible with lower limit d and upper limit d' (where $d < d'$) refers to the situation in which the insurer pays 0 if the loss is below d , the insurer pays the full loss if the loss amount is above d' , and the deductible amount reduces linearly from d to 0 as the loss increases from d to d' ;

(ii) **Policy limit**: A **policy limit of amount u** indicates that the insurer will pay a maximum amount of u on a claim.

A variation on the notion of policy limit is the **insurance cap**. An insurance cap specifies a maximum claim amount, say m , that would be paid if a loss occurs on the policy, so that the insurer pays the claim up to a maximum amount of m . If there is no deductible, this is the same as a policy limit, but if there is a deductible of d , then the maximum amount paid by the insurer is $m = u - d$. In this case, the policy limit of amount u is the same as an insurance cap of amount $u - d$

(iii) **Proportional insurance**: Proportional insurance specifies a fraction α ($0 < \alpha < 1$), and if a loss of amount X occurs, the insurer pays the policyholder αX the specified fraction of the full loss.

Reinsurance: In order to limit the exposure to catastrophic claims that can occur, insurers often set up reinsurance arrangements with other insurers. The basic forms of reinsurance are very similar algebraically to the partial insurances on individual policies described above, **but they apply to the insurer's aggregate claim random variable S** . The claims paid by the **ceding insurer** (the insurer who purchases the reinsurance) are referred to as **retained claims**.

(i) **Stop-loss reinsurance**: A stop-loss reinsurance specifies a deductible amount d . If the aggregate claim S is less than d then the reinsurer pays nothing, but if the aggregate claim is greater than d then the reinsurer pays the aggregate claim in excess of d .

(ii) **Reinsurance cap:** A reinsurance cap specifies a maximum amount paid by the reinsurer, say m .

(iii) **Proportional reinsurance:** Proportional reinsurance specifies a fraction α ($0 < \alpha < 1$), and if aggregate claims of amount S occur, the reinsurer pays αS and the ceding insurer pays $(1 - \alpha)S$.

Basic Operations on Sets

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/or, intersections/and and complements/not), 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

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

- The set of all integers

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

- The set of all rational numbers

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

- The set \mathbb{R} of all real numbers.
- The set of all complex numbers

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$$

where $i = \sqrt{-1}$.

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 either one (but not both) of the following holds:

- x belongs to A and we write $x \in A$.
- x does not belong to A , and in this case we write $x \notin A$.

Example 1.1

Which of the following is a well-defined set.

- (a) The collection of good books.
- (b) The collection of left-handed individuals in Russellville.

Solution.

- (a) The collection of good books is not a well-defined set since the answer to the question “Is *My Life* a good book?” 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 to represent 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 $A = \{x \mid x \text{ is an integer satisfying } x^2 - 4 = 0\}$.

We define the **empty** set, denoted by \emptyset , to be the set with no elements. A set which is not empty is called a **nonempty** set.

Example 1.2

List the elements of the following sets.

- (a) $\{x \mid x \text{ is a real number such that } x^2 = 1\}$.
- (b) $\{x \mid x \text{ is an integer such that } x^2 - 3 = 0\}$.

Solution.

- (a) $\{-1, 1\}$.
- (b) Since the only solutions to the given equation are $-\sqrt{3}$ and $\sqrt{3}$ and both are not integers then 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) $\{a, e, i, o, u\}$.
- (b) $\{1, 3, 5, 7, 9\}$.

Solution.

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

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 set 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 infinite set, we write $n(A) = \infty$. For example, $n(\mathbb{N}) = \infty$.

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 then the two sets are said to have the same cardinality, i.e. $n(A) = n(B)$.

Example 1.5

What is the cardinality of each of the following sets?

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

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

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 1.6

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 1.7

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

Solution.

The Venn diagram is given in Figure 1.1 ■

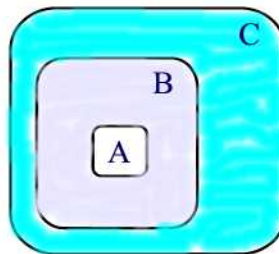


Figure 1.1

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 1.8

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

Solution.

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \blacksquare$$

Example 1.9

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 \blacksquare

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 1.10

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

Solution.

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

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 $P(1), P(2), \dots, P(n)$ are true.
 (iii) (Induction step) Show that $P(n+1)$ is true.

Example 1.11

- (a) Use induction to show that if $n(A) = n$ then $n(\mathcal{P}(A)) = 2^n$.
(b) If $\mathcal{P}(A)$ 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 $\mathcal{P}(A) = \{\emptyset\}$. Thus $n(\mathcal{P}(A)) = 1 = 2^0$. As induction hypothesis, suppose that if $n(A) = n$ then $n(\mathcal{P}(A)) = 2^n$. Let $B = \{a_1, a_2, \dots, a_n, a_{n+1}\}$. Then $\mathcal{P}(B)$ consists of all subsets of $\{a_1, a_2, \dots, a_n\}$ together with all subsets of $\{a_1, a_2, \dots, a_n\}$ with the element a_{n+1} added to them. Hence, $n(\mathcal{P}(B)) = 2^n + 2^n = 2 \cdot 2^n = 2^{n+1}$.

(b) Since $n(\mathcal{P}(A)) = 256 = 2^8$ then $n(A) = 8$ ■

Example 1.12

Use induction to show $\sum_{i=1}^n (2i - 1) = n^2$, $n \geq 1$.

Solution.

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

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 divisors: 1 and the number itself.

Problem 1.2

Consider the random experiment of tossing a coin three times.

- Let S be the collection of all outcomes of this experiment. List the elements of S . Use H for head and T for tail.
- Let E be the subset of S with more than one tail. List the elements of E .
- 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. Let E be the event that the hand contains 5 aces. List the elements of E .

Problem 1.5

Prove the following properties:

- Reflexive Property: $A \subseteq A$.
- Antisymmetric Property: If $A \subseteq B$ and $B \subseteq A$ then $A = B$.
- Transitive Property: If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.

Problem 1.6

Prove by using mathematical induction that

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

Problem 1.7

Prove by using mathematical induction that

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

Problem 1.8

Use induction to show that $(1 + x)^n \geq 1 + nx$ for all $n \geq 1$, where $x > -1$.

Problem 1.9

A caterer prepared 60 beef tacos for a birthday party. Among these tacos, he made 45 with tomatoes, 30 with both tomatoes and onions, and 5 with neither tomatoes nor onions. Using a Venn diagram, how many tacos did he make with

- (a) tomatoes or onions?
- (b) onions?
- (c) onions but not tomatoes?

Problem 1.10

A dormitory of college freshmen has 110 students. Among these students,

- 75 are taking English,
- 52 are taking history,
- 50 are taking math,
- 33 are taking English and history,
- 30 are taking English and math,
- 22 are taking history and math,
- 13 are taking English, history, and math.

How many students are taking

- (a) English and history, but not math,
- (b) neither English, history, nor math,
- (c) math, but neither English nor history,
- (d) English, but not history,
- (e) only one of the three subjects,
- (f) exactly two of the three subjects.

Problem 1.11

An experiment consists of the following two stages: (1) first a fair die is rolled (2) if the number appearing is even, then a fair coin is tossed; if the number appearing is odd, then the die 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. List the elements of S and then find the cardinality of S .

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

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

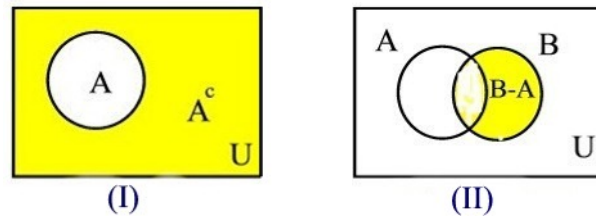


Figure 2.1

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))

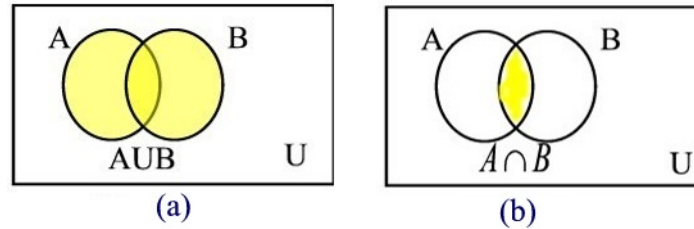


Figure 2.2

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

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

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

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

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:

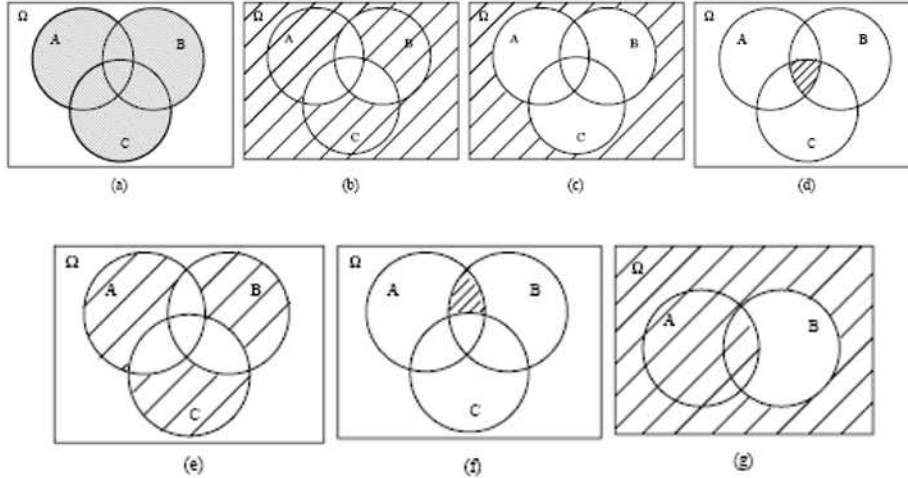
- at least one of the events A, B, C occurs;
- at most one of the events A, B, C occurs;
- none of the events A, B, C occurs;
- all three events A, B, C occur;
- exactly one of the events A, B, C occurs;
- events A and B occur, but not C ;
- either event A occurs or, if not, then B also does not occur.

In each case draw the corresponding Venn diagram.

Solution.

- $A \cup B \cup C$
- $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C) \cup (A^c \cap B^c \cap C^c)$
- $(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$
- $A \cap B \cap C$
- $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$
- $A \cap B \cap C^c$

(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
- (f) if A occurs, then B 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 intersect.

Solution.

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

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

Each team in a basketball league plays 20 games in one tournament. Event A is the event that Team 1 wins 15 or more games in the tournament. Event B is the event that Team 1 wins less than 10 games. Event C is the event that Team 1 wins between 8 to 16 games. Of course, Team 1 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 Team 1 wins 15 or more games or wins 9 or less games. $A \cap B$ is the empty set, since Team 1 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 Team 1 wins at least 8 games. $A \cap C$ is the event that Team 1 wins 15 or 16 games.

(c) $B \cup C$ is the event that Team 1 wins at most 16 games. $B \cap C$ is the event that Team 1 wins 8 or 9 games.

(d) A^c is the event that Team 1 wins 14 or fewer games. B^c is the event that Team 1 wins 10 or more games. C^c is the event that Team 1 wins fewer than 8 or more than 16 games ■

Given the sets A_1, A_2, \dots , we define

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

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 2.16 ■

Remark 2.2

Note that since \cap and \cup are commutative operations then $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ 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 then

- (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

$$(\cup_{n=1}^{\infty} A_n)^c = \cap_{n=1}^{\infty} A_n^c$$

and

$$(\cap_{n=1}^{\infty} A_n)^c = \cup_{n=1}^{\infty} A_n^c$$

Example 2.9

Let U be the set of people solicited for a contribution to a charity. All the people in U were given a chance to watch a video and to read a booklet. Let V be the set of people who watched the video, B the set of people who read the booklet, C the set of people who made a contribution.

(a) Describe with set notation: "The set of people who did not see the video or read the booklet but who still made a contribution"

(b) Rewrite your answer using De Morgan's law 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 make a contribution ■

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, \dots 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 3, B the event the total is even, and C the event the total is a multiple of 7. Show that A, B, C are pairwise disjoint.

Solution.

We have

$$A = \{(1, 2), (2, 1)\}$$

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

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

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

A total of 35 programmers interviewed for a job; 25 knew FORTRAN, 28 knew PASCAL, and 2 knew neither languages. How many knew both languages?

Solution.

Let F be the group of programmers that knew FORTRAN, P those who knew PASCAL. Then $F \cap P$ is the group of programmers who knew both languages. By the Inclusion-Exclusion Principle we have $n(F \cup P) = n(F) + n(P) - n(F \cap P)$. That is, $33 = 25 + 28 - n(F \cap P)$. Solving for $n(F \cap P)$ we find $n(F \cap P) = 20$ ■

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, \dots, A_n the Cartesian product of these sets is the set

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

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 $S = S_1 \times S_2 \times \dots \times S_n$ 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 $A = \{a_1, a_2, \dots, a_n\}$ and $B = \{b_1, b_2, \dots, b_m\}$. Then

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

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

Example 2.15

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

Solution.

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

Problems

Problem 2.1

Let A and B be any two sets. Use Venn diagram 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, 3000 are young, 4600 are male, and 7000 are married. The policyholders can also be classified as 1320 young males, 3010 married males, and 1400 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 the chewing gum preferences of baseball players, it was found that

- 22 like fruit.
- 25 like spearmint.
- 39 like grape.
- 9 like spearmint and fruit.
- 17 like fruit and grape.
- 20 like spearmint and grape.
- 6 like all flavors.
- 4 like none.

How many players 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

In a class of students undergoing a computer course the following were observed. Out of a total of 50 students:

- 30 know PASCAL,
- 18 know FORTRAN,
- 26 know COBOL,
- 9 know both PASCAL and FORTRAN,
- 16 know both Pascal and COBOL,
- 8 know both FORTRAN and COBOL,
- 47 know at least one of the three languages.

- (a) How many students know none of these languages ?
(b) How many students know all three languages ?

Problem 2.14

Mr. Brown raises chickens. Each can be described as thin or fat, brown or red, hen or rooster. Four are thin brown hens, 17 are hens, 14 are thin chickens, 4 are thin hens, 11 are thin brown chickens, 5 are brown hens, 3 are fat red roosters, 17 are thin or brown chickens. How many chickens does Mr. Brown have?

Problem 2.15 ‡

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 2.16

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 2.17

Translate the following verbal description of events into set theoretic notation. For example, “ A or B occur, 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 occur.
- (d) Neither A nor B occur.

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.

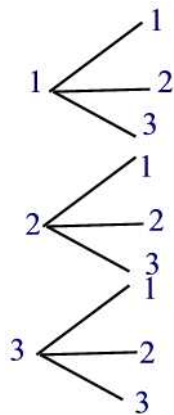
3 The Fundamental Principle of Counting

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

A lottery allows you to select a two-digit number. Each digit may be either 1, 2 or 3. Use a tree diagram to show all the possible outcomes and tell how many different numbers can be selected.

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** 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, \dots , and for each of these the k^{th} can be made in n_k ways, then the whole choice can be made in $n_1 \cdot n_2 \cdot \dots \cdot n_k$ ways.

Proof.

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

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

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

Basis of Induction

This is just Theorem 2.4.

Induction Hypothesis

Suppose

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

Induction Conclusion

We must show

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

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

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

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

Example 3.2

In designing a study of the effectiveness of migraine medicines, 3 factors were considered:

- (i) Medicine (A,B,C,D, Placebo)
- (ii) Dosage Level (Low, Medium, High)
- (iii) Dosage Frequency (1,2,3,4 times/day)

In how many possible ways can a migraine patient be given medicine?

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 migraine patient can be given medicine is $n_1 \cdot n_2 \cdot n_3 = 5 \cdot 3 \cdot 4 = 60$ different ways ■

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 ■

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 ■

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 $\{0, 2, \dots, 9\}$. 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 ■

Problems

Problem 3.1

If each of the 10 digits 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.
- (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.

Problem 3.2

- (a) If eight horses 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 horses are Lucky one, Lucky Two, and Lucky Three, in how many possible orders can they finish?

Problem 3.3

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

Problem 3.4

A club has 10 members. In how many ways can the club choose a president and vice-president if everyone is eligible?

Problem 3.5

In a medical study, patients are classified according to whether they have blood type A, B, AB, or O, 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, or bus, in how many different ways can such a trip be arranged?

Problem 3.7

If twenty paintings are entered in an art show, 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 52 members of a labor union 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 movies can be ranked first, second, third, and fourth according to their attendance figures for the first six months.

Problem 3.10

How many ways are there to seat 10 people, consisting of 5 couples, in a row of seats (10 seats wide) if all couples are to get adjacent seats?

4 Permutations and Combinations

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, \dots , the 8th step is the possibility of a horse to finish 8th 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

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

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) $\frac{10!}{7!} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 10 \cdot 9 \cdot 8 = 720$ ■

Using factorials we see that the number of permutations of n objects is $n!$.

Example 4.2

There are $6!$ permutations of the 6 letters of the word “square.” In how many of them is r the second letter?

Solution.

Let r be the second letter. Then there are 5 ways to fill the first spot, 4 ways to fill the third, 3 to fill the fourth, and so on. There are $5!$ 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 $P(n, k)$. 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 $P(n, k)$ is given next.

Theorem 4.1

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

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

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, $P(n, k) = n(n - 1) \cdots (n - k + 1) = \frac{n(n-1) \cdots (n-k+1)(n-k)!}{(n-k)!} = \frac{n!}{(n-k)!}$ ■

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 $P(26, 3)$ ways. The second step is to select the digits and this can be done in $P(10, 4)$ ways. Thus, by the Fundamental Principle of Counting there are $P(26, 3) \cdot P(10, 4) = 78,624,000$ license plates ■

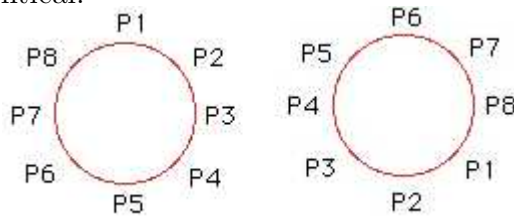
Example 4.5

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

Solution.

$$P(10, 5) = \frac{10!}{(10-5)!} = 30,240 \text{ zip codes } \blacksquare$$

Circular permutations are ordered arrangements of objects in a circle. While order is still considered, a circular permutation is not considered to be distinct from another unless at least one object is preceded or succeeded by a different object in both permutations. Thus, the following permutations are considered identical.



In the set of objects, one object can be fixed, and the other objects can be arranged in different permutations. Thus, the number of permutations of n distinct objects that are arranged in a circle is $(n - 1)!$.

Example 4.6

In how many ways can you seat 6 persons at a circular dinner table.

Solution.

There are $(6 - 1)! = 5! = 120$ ways to seat 6 persons at a circular dinner table. \blacksquare

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 $C(n, k)$ and is read “ n choose k ”. The formula for $C(n, k)$ is given next.

Theorem 4.2

If $C(n, k)$ denotes the number of ways in which k objects can be selected from a set of n distinct objects then

$$C(n, k) = \frac{P(n, k)}{k!} = \frac{n!}{k!(n-k)!}.$$

Proof.

Since the number of groups of k elements out of n elements is $C(n, k)$ and each group can be arranged in $k!$ ways then $P(n, k) = k!C(n, k)$. It follows that

$$C(n, k) = \frac{P(n, k)}{k!} = \frac{n!}{k!(n-k)!} \blacksquare$$

An alternative notation for $C(n, k)$ is $\binom{n}{k}$. We define $C(n, k) = 0$ if $k < 0$ or $k > n$.

Example 4.7

From a group of 5 women and 7 men, how many different committees consisting of 2 women and 3 men can be formed? What if 2 of the men are feuding and refuse to serve on the committee together?

Solution.

There are $C(5, 2)C(7, 3) = 350$ possible committees consisting of 2 women and 3 men. Now, if we suppose that 2 men are feuding and refuse to serve together then the number of committees that do not include the two men is $C(7, 3) - C(2, 2)C(5, 1) = 30$ possible groups. Because there are still $C(5, 2) = 10$ possible ways to choose the 2 women, it follows that there are $30 \cdot 10 = 300$ possible committees \blacksquare

The next theorem discusses some of the properties of combinations.

Theorem 4.3

Suppose that n and k are whole numbers with $0 \leq k \leq n$. Then

- (a) $C(n, 0) = C(n, n) = 1$ and $C(n, 1) = C(n, n-1) = n$.
- (b) Symmetry property: $C(n, k) = C(n, n-k)$.
- (c) Pascal's identity: $C(n+1, k) = C(n, k-1) + C(n, k)$.

Proof.

- a. From the formula of $C(\cdot, \cdot)$ we have $C(n, 0) = \frac{n!}{0!(n-0)!} = 1$ and $C(n, n) =$

$\frac{n!}{n!(n-n)!} = 1$. Similarly, $C(n, 1) = \frac{n!}{1!(n-1)!} = n$ and $C(n, n-1) = \frac{n!}{(n-1)!} = n$.

b. Indeed, we have $C(n, n-k) = \frac{n!}{(n-k)!(n-n+k)!} = \frac{n!}{k!(n-k)!} = C(n, k)$.

c.

$$\begin{aligned} C(n, k-1) + C(n, k) &= \frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!} \\ &= \frac{n!k}{k!(n-k+1)!} + \frac{n!(n-k+1)}{k!(n-k+1)!} \\ &= \frac{n!}{k!(n-k+1)!} (k+n-k+1) \\ &= \frac{(n+1)!}{k!(n+1-k)!} = C(n+1, k) \blacksquare \end{aligned}$$

Pascal's identity allows one to construct the so-called Pascal's triangle (for $n = 10$) as shown in Figure 4.1.

										n
										0
										1
										2
										3
										4
										5
										6
										7
										8
										9
										10

Figure 4.1

Example 4.8

The Chess Club 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 regional tournament with no regard to order?

Solution.

(a) $P(6, 6) = 6! = 720$ different ways

- (b) $P(6, 2) = 30$ ways
 (c) $C(6, 3) = 20$ different ways ■

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 4.4 (*Binomial Theorem*)

Let x and y be variables, and let n be a non-negative integer. Then

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

where $C(n, k)$ will be called the **binomial coefficient**.

Proof.

The proof is by induction on n .

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

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

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

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

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

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

Indeed, we have

$$\begin{aligned}
 (x + y)^{n+1} &= (x + y)(x + y)^n = x(x + y)^n + y(x + y)^n \\
 &= x \sum_{k=0}^n C(n, k)x^{n-k}y^k + y \sum_{k=0}^n C(n, k)x^{n-k}y^k \\
 &= \sum_{k=0}^n C(n, k)x^{n-k+1}y^k + \sum_{k=0}^n C(n, k)x^{n-k}y^{k+1} \\
 &= [C(n, 0)x^{n+1} + C(n, 1)x^n y + C(n, 2)x^{n-1}y^2 + \cdots + C(n, n)xy^n] \\
 &\quad + [C(n, 0)x^n y + C(n, 1)x^{n-1}y^2 + \cdots + C(n, n-1)xy^n + C(n, n)y^{n+1}] \\
 &= C(n + 1, 0)x^{n+1} + [C(n, 1) + C(n, 0)]x^n y + \cdots + \\
 &\quad [C(n, n) + C(n, n-1)]xy^n + C(n + 1, n + 1)y^{n+1} \\
 &= C(n + 1, 0)x^{n+1} + C(n + 1, 1)x^n y + C(n + 1, 2)x^{n-1}y^2 + \cdots \\
 &\quad + C(n + 1, n)xy^n + C(n + 1, n + 1)y^{n+1} \\
 &= \sum_{k=0}^{n+1} C(n + 1, k)x^{n-k+1}y^k. \blacksquare
 \end{aligned}$$

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 4.9

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 4.10

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

Solution.

Since there are $C(n, k)$ subsets of k elements with $0 \leq k \leq n$, the total number of subsets of a set of n elements is

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

Problems

Problem 4.1

Find m and n so that $P(m, n) = \frac{9!}{6!}$

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 no repetitions of letters are permitted, 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 combination lock has 40 numbers on it.

- (a) How many different three-number combinations can be made?
- (b) How many different combinations are there if the three numbers are different?

Problem 4.5

(a) Miss Murphy wants to seat 12 of her students in a row for a class picture. How many different seating arrangements are there?

(b) Seven of Miss Murphy's students are girls and 5 are boys. In how many different ways can she seat the 7 girls together on the left, and then the 5 boys 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 officers, a president and a treasurer, be chosen?

Problem 4.8

(a) A baseball team has nine players. Find the number of ways the manager can arrange the batting order.

(b) 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.

Problem 4.9

Find m and n so that $C(m, n) = 13$

Problem 4.10

The Library of Science Book Club offers three books from a list of 42. If you circle three choices from a list of 42 numbers on a postcard, how many possible choices are there?

Problem 4.11

At the beginning of the second quarter of a mathematics class for elementary school teachers, each of the class's 25 students shook hands with each of the other students exactly once. How many handshakes took place?

Problem 4.12

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

Problem 4.13

A consumer group plans to select 2 televisions from a shipment of 8 to check the picture quality. In how many ways can they choose 2 televisions?

Problem 4.14

A school has 30 teachers. In how many ways can the school choose 3 people to attend a national meeting?

Problem 4.15

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

Problem 4.16

How many different 12-person juries can be chosen from a pool of 20 jurors?

Problem 4.17

A jeweller has 15 different sized pearls to string on a circular band. In how many ways can this be done?

Problem 4.18

Four teachers and four students are seated in a circular discussion group. Find the number of ways this can be done if teachers and students must be seated alternately.

5 Permutations and Combinations with Indistinguishable Objects

So far we have looked at permutations and combinations of objects where each object was distinct (distinguishable). Now we will consider allowing multiple copies of the same (identical/indistinguishable) item.

Permutations with Repetitions (i.e. Indistinguishable Objects)

We have seen how to calculate basic permutation problems where each element can be used one time at most. In the following discussion, we will see what happens when elements can be used repeatedly. As an example, consider the following problem: How many different letter arrangements can be formed using the letters in DECEIVED? Note that the letters in DECEIVED are not all distinguishable since it contains repeated letters such as E and D . Thus, interchanging the second and the fourth letters gives a result indistinguishable from interchanging the second and the seventh letters. The number of different rearrangement of the letters in DECEIVED can be found by applying the following theorem.

Theorem 5.1

Given n objects of which n_1 indistinguishable objects of type 1, n_2 indistinguishable objects of type 2, \dots , n_k indistinguishable objects of type k where $n_1 + n_2 + \dots + n_k = n$. Then the number of distinguishable permutations of n objects is given by:

$$\frac{n!}{n_1!n_2!\cdots n_k!}.$$

Proof.

The task of forming a permutation can be divided into k subtasks: For the first kind we have n slots and n_1 objects to place such that the order is unimportant. There are $C(n, n_1)$ different ways. For the second kind we have $n - n_1$ slots and n_2 objects to place such that the order is unimportant. There are $C(n - n_1, n_2)$ ways. For the k^{th} kind we have $n - n_1 - n_2 - \dots - n_{k-1}$ slots and n_k objects to place such that the order is unimportant. There are $C(n - n_1 - n_2 - \dots - n_{k-1}, n_k)$ different ways.

Thus, applying the Fundamental Principle of Counting we find that the number of distinguishable permutations is

$$C(n, n_1) \cdot C(n - n_1, n_2) \cdots C(n - n_1 - \dots - n_{k-1}, n_k) = \frac{n!}{n_1!n_2!\cdots n_k!} \blacksquare$$

It follows that the number of different rearrangements of the letters in DECEIVED is $\frac{8!}{2!3!} = 3360$.

Example 5.1

In how many ways can you arrange 2 red, 3 green, and 5 blue balls in a row?

Solution.

There are $\frac{(2+3+5)!}{2!3!5!} = 2520$ different ways ■

Example 5.2

How many strings can be made using 4 A's, 3 B's, 7 C's and 1 D?

Solution.

Applying the previous theorem with $n = 4 + 3 + 7 + 1 = 15$, $n_1 = 4$, $n_2 = 3$, $n_3 = 7$, and $n_4 = 1$ we find that the total number of different such strings is

$$\frac{15!}{7!4!3!1!} = 1,801,800 \blacksquare$$

Distributing Distinct Objects into Distinguishable Boxes

We consider the following problem: In how many ways can you distribute n distinct objects into k different boxes so that there will be n_1 objects in Box 1, n_2 in Box 2, \dots , n_k in Box k ? where $n_1 + n_2 + \dots + n_k = n$.

The answer to this question is provided by the following theorem.

Theorem 5.2

The number of ways to distribute n distinct objects into k distinguishable boxes so that there are n_i objects placed into box i , $1 \leq i \leq k$ is:

$$\frac{n!}{n_1!n_2! \cdots n_k!}$$

Proof.

The proof is a repetition of the proof of Theorem 5.1: There are $C(n, n_1)$ possible choices for the first box; for each choice of the first box there are $C(n - n_1, n_2)$ possible choices for the second box; for each possible choice of the first two boxes there are $C(n - n_1 - n_2)$ possible choices for the third box and so on. According to Theorem 3.1, there are

$$C(n, n_1)C(n - n_1, n_2)C(n - n_1 - n_2, n_3) \cdots C(n - n_1 - n_2 - \cdots - n_{k-1}, n_k) = \frac{n!}{n_1!n_2! \cdots n_k!}$$

different ways ■

Example 5.3

How many ways are there of assigning 10 police officers to 3 different tasks:

- Patrol, of which there must be 5.
- Office, of which there must be 2.
- Reserve, of which there will be 3.

Solution.

There are $\frac{10!}{5!2!3!} = 2520$ different ways ■

A correlation between Theorem 5.1 and Theorem 5.2 is shown next.

Example 5.4

Solve Example 5.2 using Theorem 5.2.

Solution.

Consider the string of length 15 composed of A's, B's, C's, and D's from Example 5.2. Instead of placing the letters into the 15 unique positions, we place the 15 positions into the 4 boxes which represent the 4 letters. Now apply Theorem 5.2 ■

Example 5.5

Consider a group of 10 different people. Two committees are to be formed, one of four people and one of three people. (No person can serve on both committees.) In how many ways can the committees be formed.

Solution.

The three remaining people not chosen can be viewed as a third “committee”, so the number of choices is just the multinomial coefficient

$$\binom{10}{4, 3, 3} = \frac{10!}{4!3!3!} = 4200 \blacksquare$$

Distributing n Identical Objects into k Boxes

Consider the following problem: How many ways can you put 4 identical balls into 2 boxes? We will draw a picture in order to understand the solution to this problem. Imagine that the 4 identical items are drawn as stars: $\{****\}$.

If we draw a vertical bar somewhere among these 4 stars, this can represent a unique assignment of the balls to the boxes. For example:

	<i>BOX 1</i>	<i>BOX 2</i>
* * * *	0	4
* * * *	1	3
* * * *	2	2
* * * *	3	1
* * * *	4	0

Since there is a correspondence between a star/bar picture and assignments of balls to boxes, we can count the star/bar pictures and get an answer to our balls-in-boxes problem. So the question of finding the number of ways to put four identical balls into 2 boxes is the same as the number of ways of arranging four identical stars and one vertical bar which is $\binom{5}{4} = \binom{5}{1} = 5$. This is what one gets when replacing $n = 4$ and $k = 2$ in the following theorem.

Theorem 5.3

The number of ways n identical (i.e. indistinguishable) objects can be distributed into k boxes is

$$\binom{n+k-1}{k-1} = \binom{n+k-1}{n}.$$

Proof.

Imagine the n identical objects as n stars. Draw $k-1$ vertical bars somewhere among these n stars. This can represent a unique assignment of the balls to the boxes. Hence, there is a correspondence between a star/bar picture and assignments of balls to boxes. So how many ways can you arrange n identical dots and $k-1$ vertical bar? The answer is given by

$$\binom{n+k-1}{k-1} = \binom{n+k-1}{n} \blacksquare$$

It follows that there are $C(n+k-1, k-1)$ ways of placing n identical objects into k distinct boxes.

Example 5.6

How many ways can we place 7 identical balls into 8 separate (but distinguishable) boxes?

Solution.

We have $n = 7$, $k = 8$ so there are $C(8 + 7 - 1, 7)$ different ways ■

Example 5.7

An ice cream store sells 21 flavors. Fred goes to the store and buys 5 quarts of ice cream. How many choices does he have?

Solution.

Fred wants to choose 5 things from 21 things, where order doesn't matter and repetition is allowed. The number of possibilities is $C(21 + 5 - 1, 5) = 53130$ ■

Remark 5.4

Note that Theorem 5.3 gives the number of vectors (n_1, n_2, \dots, n_k) , where n_i is a nonnegative integer, such that

$$n_1 + n_2 + \dots + n_k = n$$

where n_i is the number of objects in box i , $1 \leq i \leq k$.

Example 5.8

How many solutions to $x_1 + x_2 + x_3 = 11$ with x_i nonnegative integers.

Solution.

This is like throwing 11 objects in 3 boxes. The objects inside each box are indistinguishable. Let x_i be the number of objects in box i where $i = 1, 2, 3$. Then the number of solutions to the given equation is

$$\binom{11 + 3 - 1}{11} = 78. \blacksquare$$

Example 5.9

How many solutions are there to the equation

$$n_1 + n_2 + n_3 + n_4 = 21$$

where $n_1 \geq 2$, $n_2 \geq 3$, $n_3 \geq 4$, and $n_4 \geq 5$?

Solution.

We can rewrite the given equation in the form

$$(n_1 - 2) + (n_2 - 3) + (n_3 - 4) + (n_4 - 5) = 7$$

or

$$m_1 + m_2 + m_3 + m_4 = 7$$

where the m_i s are positive. The number of solutions is $C(7+4-1, 7) = 120$ ■

Multinomial Theorem

We end this section by extending the binomial theorem to the multinomial theorem when the original expression has more than two variables.

Theorem 5.4

For any variables x_1, x_2, \dots, x_r and any nonnegative integer n we have

$$(x_1 + x_2 + \dots + x_r)^n = \sum_{\substack{(n_1, n_2, \dots, n_r) \\ n_1 + n_2 + \dots + n_r = n}} \frac{n!}{n_1!n_2!\dots n_r!} x_1^{n_1} x_2^{n_2} \dots x_r^{n_r}.$$

Proof.

We provide a combinatorial proof. A generic term in the expansion of $(x_1 + x_2 + \dots + x_r)^n$ will take the form $Cx_1^{n_1}x_2^{n_2}\dots x_r^{n_r}$ with $n_1 + n_2 + \dots + n_r = n$. We seek a formula for the value of C .

Obtaining a term of the form $x_1^{n_1}x_2^{n_2}\dots x_r^{n_r}$ involves choosing n_1 terms of n terms of our expansion to be occupied by x_1 , then choosing n_2 out of the remaining $n - n_1$ terms for x_2 , choosing n_3 of the remaining $n - n_1 - n_2$ terms for x_3 and so on. Thus, the number of times $x_1^{n_1}x_2^{n_2}\dots x_r^{n_r}$ occur in the expansion is

$$\begin{aligned} C &= C(n, n_1)C(n - n_1, n_2)C(n - n_1 - n_2, n_3) \dots C(n - n_1 - \dots - n_{r-1}, n_r) \\ &= \frac{n!}{n_1!n_2!\dots n_r!} \quad \blacksquare \end{aligned}$$

It follows from Remark 5.4 that the number of coefficients in the multinomial expansion is given by $C(n + r - 1, r - 1)$.

Example 5.10

What is the coefficient of $x^3y^6z^{12}$ in the expansion of $(x + 2y^2 + 4z^3)^{10}$?

Solution.

If we expand using the trinomial theorem (three term version of the multinomial theorem), then we have

$$(x + 2y^2 + 4z^3)^{10} = \binom{10}{n_1, n_2, n_3} x^{n_1} (2y^2)^{n_2} (4z^3)^{n_3}.$$

where the sum is taken over all triples of numbers (n_1, n_2, n_3) such that $n_i \geq 0$ and $n_1 + n_2 + n_3 = 10$. If we want to find the term containing $x^2y^6z^{12}$, we must choose $n_1 = 3, n_2 = 3$, and $n_3 = 4$, and this term looks like

$$\binom{10}{3, 3, 4} x^3(2y^2)^3(4z^3)^4 = \frac{10!}{3!3!4!} 2^3 4^3 x^3 y^6 z^{12}.$$

The coefficient of this term is $\frac{10!}{3!3!4!} 2^3 4^3$ ■

Example 5.11

How many distinct 6-letter “words” can be formed from 3 A’s, 2 B’s, and a C? Express your answer as a number.

Solution.

There are

$$\binom{6}{3, 2, 1} = \frac{6!}{3!2!1!} = 60 \blacksquare$$

Problems

Problem 5.1

How many rearrangements of the word MATHEMATICAL are there?

Problem 5.2

Suppose that a club with 20 members plans to form 3 distinct committees with 6, 5, and 4 members, respectively. In how many ways can this be done?

Problem 5.3

A deck of cards contains 13 clubs, 13 diamonds, 13 hearts and 13 spades, with 52 cards in all. In a game of bridge, 13 cards are dealt to each of 4 persons. In how many ways the 52 cards can be dealt evenly among 4 players?

Problem 5.4

A book publisher has 3000 copies of a discrete mathematics book. How many ways are there to store these books in their three warehouses if the copies of the book are indistinguishable?

Problem 5.5

A chess tournament has 10 competitors of which 4 are from Russia, 3 from the United States, 2 from Great Britain, and 1 from Brazil. If the tournament result lists just the nationalities of the players in the order in which they placed, how many outcomes are possible?

Problem 5.6

How many different signals, each consisting of 9 flags hung in a line, can be made from a set of 4 white flags, 3 red flags, and 2 blue flags if all flags of the same color are identical?

Problem 5.7

An investor has \$20,000 to invest among 4 possible investments. Each investment must be a unit of \$1,000. If the total \$20,000 is to be invested, how many different investment strategies are possible? What if not all the money need to be invested?

Problem 5.8

How many solutions does the equation $x_1 + x_2 + x_3 = 5$ have such that x_1, x_2, x_3 are non negative integers?

Problem 5.9

- (a) How many distinct terms occur in the expansion of $(x + y + z)^6$?
(b) What is the coefficient of xy^2z^3 ?
(c) What is the sum of the coefficients of all the terms?

Problem 5.10

How many solutions are there to $x_1 + x_2 + x_3 + x_4 + x_5 = 50$ for integers $x_i \geq 1$?

Problem 5.11

A probability teacher buys 36 plain doughnuts for his class of 25. In how many ways can the doughnuts be distributed? (Note: Not all students need to have received a doughnut.)

Problem 5.12

- (a) How many solutions exist to the equation $x_1 + x_2 + x_3 = 15$, where x_1, x_2 , and x_3 have to be non-negative integers? Simplify your answer as much as possible. [Note: the solution $x_1 = 12, x_2 = 2, x_3 = 1$ is not the same as $x_1 = 1, x_2 = 2, x_3 = 12$]
(b) How many solutions exist to the equation $x_1x_2x_3 = 3^62^{13}$, where x_1, x_2 and x_3 have to be positive integers? Hint: Let $x_i = 2^{a_i}3^{b_i}$, $i = 1, 2, 3$.

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 Basic Definitions and Axioms of Probability

An **experiment** is any situation 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.

By an **outcome** or **simple event** we mean any result of the experiment. For example, the experiment of rolling a die yields six outcomes, namely, the outcomes 1, 2, 3, 4, 5, and 6.

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 simple events $\{1, 3, 5\}$.

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 simple events:

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

- (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 $P(E)$, the probability of the event E , is defined by

$$P(E) = \lim_{n \rightarrow \infty} \frac{n(E)}{n}.$$

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 $P(E)$. This result is a theorem called the law of large numbers which we will discuss in Section 39.1.

The function P satisfies the following axioms, known as **Kolmogorov axioms**:

Axiom 1: For any event E , $0 \leq P(E) \leq 1$.

Axiom 2: $P(S) = 1$.

Axiom 3: For any sequence of mutually exclusive events $\{E_n\}_{n \geq 1}$, that is $E_i \cap E_j = \emptyset$ for $i \neq j$, we have

$$P(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} P(E_n). \text{ (Countable Additivity)}$$

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

$$P(\cup_{k=1}^n E_k) = \sum_{k=1}^n P(E_k).$$

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

Example 6.2

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

Solution.

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

Example 6.3

If, for a given experiment, O_1, O_2, O_3, \dots is an infinite sequence of outcomes, verify that

$$P(O_i) = \left(\frac{1}{2}\right)^i, \quad i = 1, 2, 3, \dots$$

is a probability measure.

Solution.

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

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

Now, if E_1, E_2, \dots is a sequence of mutually exclusive events then

$$P(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} P(O_{nj}) = \sum_{n=1}^{\infty} P(E_n)$$

where $E_n = \{O_{n1}, O_{n2}, \dots\}$. Thus, P defines a probability function ■.

Now, since $E \cup E^c = S$, $E \cap E^c = \emptyset$, and $P(S) = 1$ then

$$P(E^c) = 1 - P(E).$$

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

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

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

Since for any event E we have $\emptyset \subseteq E \subseteq S$ then $0 \leq n(E) \leq n(S)$ so that $0 \leq \frac{n(E)}{n(S)} \leq 1$. It follows that $0 \leq P(E) \leq 1$. Clearly, $P(S) = 1$. 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 .

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 $P(E) = 0$ ■

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.

Since the event of having a 3 or a 4 has two simple events $\{3, 4\}$ then the probability of rolling a 3 or a 4 is $\frac{2}{6} = \frac{1}{3}$. ■

Example 6.7

In a room containing n people, calculate the chance that at least two of them have the same birthday.

Solution.

We have

$$P(\text{Two or more have birthday match}) = 1 - P(\text{no birthday match})$$

Since each person was born on one of the 365 days in the year, there are $(365)^n$ possible outcomes (assuming no one was born in Feb 29). Now,

$$P(\text{no birthday match}) = \frac{(365)(364)\cdots(365-n+1)}{(365)^n}$$

Thus,

$$\begin{aligned} P(\text{Two or more have birthday match}) &= 1 - P(\text{no birthday match}) \\ &= 1 - \frac{(365)(364)\cdots(365-n+1)}{(365)^n} \quad \blacksquare \end{aligned}$$

Remark 6.5

It is important to keep in mind that the above 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 $S = \{\text{Red}, \text{Blue}\}$, but the outcome Blue is more likely to occur than is the outcome Red. Indeed, $P(\text{Blue}) = \frac{3}{4}$ whereas $P(\text{Red}) = \frac{1}{4}$ as supposed to $P(\text{Blue}) = P(\text{Red}) = \frac{1}{2}$

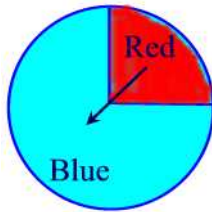


Figure 6.1

Problems

Problem 6.1

Consider the random experiment of rolling a die.

- Find the sample space of this experiment.
- Find the event of rolling the die an even number.

Problem 6.2

An experiment consists of the following two stages: (1) first a fair die is rolled (2) if the number appearing is even, then a fair coin is tossed; if the number appearing is odd, then the die 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 four-faced dice.

- Write down the sample space of this experiment.
- If E is the event "total score is at least 4", list the outcomes belonging to E^c .
- If each die is fair, find the probability that the total score is at least 6 when the two dice are thrown. What is the probability that the total score is less than 6?
- What is the probability that a double: (i.e. $\{(1, 1), (2, 2), (3, 3), (4, 4)\}$) will not be thrown?
- What is the probability that a double is not thrown nor is the score greater than 6?

Problem 6.5

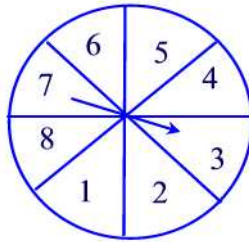
Let $S = \{1, 2, 3, \dots, 25\}$. If a number is chosen at **random**, that is, with the

same chance of being drawn as all other numbers in the set, calculate each of the following probabilities:

- The event A that an even number is drawn.
- The event B that a number less than 10 and greater than 20 is drawn.
- The event C that a number less than 26 is drawn.
- The event D that a prime number is drawn.
- The event E that a number both even and prime is drawn.

Problem 6.6

The following spinner is spun:



Find the probabilities of obtaining each of the following:

- $P(\text{factor of } 35)$
- $P(\text{multiple of } 3)$
- $P(\text{even number})$
- $P(11)$
- $P(\text{composite number})$
- $P(\text{neither prime nor composite})$

Problem 6.7

The game of bridge is played by four players: north, south, east and west. Each of these players receive 13 cards.

- What is the probability that one of the players receives all 13 spades?
- Consider a single hand of bridge. What is the probability for the hand to have 3 Aces?

Problem 6.8

A cooler contains 15 cans of Coke and 10 cans of Pepsi. Three are drawn from the cooler without replacement. What is the probability that all three are the same brand?

Problem 6.9

A coin is tossed repeatedly. What is the probability that the second head

appears at the 5th toss? (Hint: Since only the first five tosses matter, you can assume that the coin is tossed only 5 times.)

Problem 6.10

Suppose two n -sided dice are rolled. Define an appropriate probability space S and find the probabilities of the following events:

- (a) the maximum of the two numbers rolled is less than or equal to 2;
- (b) the maximum of the two numbers rolled is exactly equal to 3.

Problem 6.11

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

Problem 6.12

A fashionable club has 100 members, 30 of whom are lawyers. 25 of the members are liars, while 55 are neither lawyers nor liars.

- (a) How many of the lawyers are liars?
- (b) A liar is chosen at random. What is the probability that he is a lawyer?

Problem 6.13

An inspector selects 2 items at random from a shipment of 5 items, of which 2 are defective. She 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 shipment will not be accepted if both sampled items are defective. What is the probability she will not accept the shipment?

7 Properties of Probability

In this section we discuss some of the important properties of $P(\cdot)$.

We define the probability of nonoccurrence of an event E (called its **failure**) to be the number $P(E^c)$. Since $S = E \cup E^c$ and $E \cap E^c = \emptyset$ then $P(S) = P(E) + P(E^c)$. Thus,

$$P(E^c) = 1 - P(E).$$

Example 7.1

The probability that a college student without a flu shot will get the flu is 0.45. What is the probability that a college student without the flu shot will not get the flu?

Solution.

Our sample space consists of those students who did not get the flu shot. Let E be the set of those students without the flu shot who did get the flu. Then $P(E) = 0.45$. The probability that a student without the flu shot will not get the flu is then $P(E^c) = 1 - P(E) = 1 - 0.45 = 0.55$ ■

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 $P(A \cap B) = P(\emptyset) = 0$.

Example 7.2

Consider the sample space of rolling a die. Let A be the event of rolling an even number, B the event of rolling an odd number, and C the event of rolling a 2. Find

- $A \cup B, A \cup C,$ and $B \cup C$.
- $A \cap B, A \cap C,$ and $B \cap C$.
- Which events are mutually exclusive?

Solution.

- We have

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

(b)

$$A \cap B = \emptyset$$

$$A \cap C = \{2\}$$

$$B \cap C = \emptyset$$

(c) A and B are mutually exclusive as well as B and C ■**Example 7.3**

Let A be the event of drawing a “king” from a well-shuffled standard deck of playing cards and B the event of drawing a “ten” card. Are A and B mutually exclusive?

Solution.

Since $A = \{\text{king of diamonds, king of hearts, king of clubs, king of spades}\}$ and $B = \{\text{ten of diamonds, ten of hearts, ten of clubs, ten of spades}\}$ then A and B are mutually exclusive ■

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

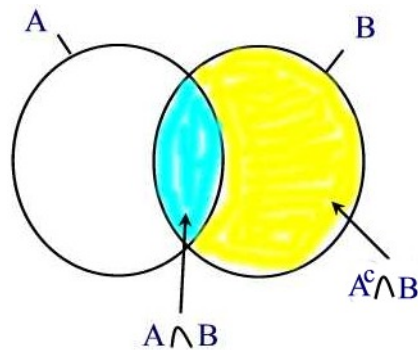
Theorem 7.1

Let A and B be two events. Then

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

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 below we see that $B = (A \cap B) \cup (A^c \cap B)$ and $A \cup B = A \cup (A^c \cap B)$.



Since $(A \cap B)$ and $(A^c \cap B)$ are mutually exclusive, by Axiom 3

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

Thus,

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

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

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

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

$$P(A \cup B) = P(A) + P(B) \quad (A, B \text{ mutually exclusive})$$

Example 7.4

A mall has two elevators. Let A be the event that the first elevator is busy, and let B be the event the second elevator is busy. Assume that $P(A) = 0.2$, $P(B) = 0.3$ and $P(A \cap B) = 0.06$. Find the probability that neither of the elevators is busy.

Solution.

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

Example 7.5

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

Solution.

Since $P(A) + P(B) = 1.5$ and $0 \leq P(A \cup B) \leq 1$ then by the previous theorem

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) \geq 1.5 - 1 = 0.5.$$

So the minimum value of $P(A \cap B)$ is 0.5 \blacksquare

Example 7.6

Suppose there's 40% chance of colder weather, 10% chance of rain and colder weather, 80% chance of rain or colder weather. Find the chance of rain.

Solution.

By the addition rule we have

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

Example 7.7

Suppose S is the set of all positive integers such that $P(s) = \frac{2}{3^s}$ for all $s \in S$. What is the probability that a number chosen at random from S will be even?

Solution.

We have

$$\begin{aligned} P(\{2, 4, 6, \dots\}) &= P(2) + P(4) + P(6) + \dots \\ &= \frac{2}{3^2} + \frac{2}{3^4} + \frac{2}{3^6} + \dots \\ &= \frac{2}{3^2} \left[1 + \frac{1}{3^2} + \frac{1}{3^4} + \dots \right] \\ &= \frac{2}{9} \left[1 + \frac{1}{9} + \frac{1}{9^2} + \frac{1}{9^3} + \dots \right] \\ &= \frac{2}{9} \cdot \frac{1}{1 - \frac{1}{9}} = \frac{1}{4} \blacksquare \end{aligned}$$

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

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

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

$$E \subseteq F \implies P(E) \leq P(F).$$

Theorem 7.2

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

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

Proof.

We have

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

Example 7.8

If a person visits his dentist, suppose that the probability that he will have his teeth cleaned is 0.44, the probability that he will have a cavity filled is 0.24, the probability that he will have a tooth extracted is 0.21, the probability that he will have cleaned and a cavity filled is 0.08, the probability that he will have his teeth cleaned and a tooth extracted is 0.11, the probability that he will have a cavity filled and a tooth extracted is 0.07, and the probability that he will have his teeth cleaned, a cavity filled, and a tooth extracted is 0.03. What is the probability that a person visiting his dentist will have at least one of these things done to him?

Solution.

Let C be the event that a person will have his teeth cleaned, F is the event that a person will have a cavity filled, and E is the event a person will have a tooth extracted. We are given $P(C) = 0.44$, $P(F) = 0.24$, $P(E) = 0.21$, $P(C \cap F) = 0.08$, $P(C \cap E) = 0.11$, $P(F \cap E) = 0.07$ and $P(C \cap F \cap E) = 0.03$. Thus,

$$P(C \cup F \cup E) = 0.44 + 0.24 + 0.21 - 0.08 - 0.11 - 0.07 + 0.03 = 0.66 \blacksquare$$

Problems

Problem 7.1

If A and B are the events that a consumer testing service will rate a given stereo system very good or good, $P(A) = 0.22$, $P(B) = 0.35$. Find

- (a) $P(A^c)$;
- (b) $P(A \cup B)$;
- (c) $P(A \cap B)$.

Problem 7.2

If the probabilities are 0.20, 0.15, and 0.03 that a student will get a failing grade in Statistics, in English, or in both, what is the probability that the student will get a failing grade in at least one of these subjects?

Problem 7.3

If A is the event “drawing an ace” from a deck of cards and B is the event “drawing a spade”. Are A and B mutually exclusive? Find $P(A \cup B)$.

Problem 7.4

A bag contains 18 colored marbles: 4 are colored red, 8 are coloured yellow and 6 are colored green. A marble is selected at random. What is the probability that the marble chosen is either red or green?

Problem 7.5

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

Problem 7.6

A golf bag contains 2 red tees, 4 blue tees, and 5 white tees.

- (a) What is the probability of the event R that a tee drawn at random is red?
- (b) What is the probability of the event “not R ” that is, that a tee drawn at random is not red?
- (c) What is the probability of the event that a tee drawn at random is either red or blue?

Problem 7.7

A fair pair of dice is rolled. Let E be 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

If events A and B are from the same sample space, and if $P(A)=0.8$ and $P(B)=0.9$, can events A and B be mutually exclusive?

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 $P(A \cup B) = 0.7$ and $P(A \cup B^c) = 0.9$. Determine $P(A)$.

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

Near a certain exit of I-17, the probabilities are 0.23 and 0.24 that a truck stopped at a roadblock will have faulty brakes or badly worn tires. Also, the probability is 0.38 that a truck stopped at the roadblock will have faulty brakes and/or badly worn tires. What is the probability that a truck stopped at this roadblock will have faulty brakes as well as badly worn tires?

Problem 7.15 †

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.

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

A quiz has 5 multiple-choice questions. Each question has 4 answer choices, of which 1 is correct and the other 3 are incorrect. Suppose that you guess on each question.

- (a) How many ways are there to answer the 5 questions?
- (b) What is the probability of getting all 5 questions right?
- (c) What is the probability of getting exactly 4 questions right and 1 wrong?
- (d) What is the probability of doing well (getting at least 4 right)?

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

$$(4)(4)(4)(4)(4) = 1024 \text{ ways}$$

(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 possibilities. Hence,

$$P(\text{all 5 right}) = \frac{1}{1024}$$

(c) The following table lists all possible responses that involve at least 4 right answers, 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

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

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

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

Example 8.2

Suppose $S = \{1, 2, 3, 4, 5, 6\}$. How many events A are there with $P(A) = \frac{1}{3}$?

Solution.

We must have $P(\{i, j\}) = \frac{1}{3}$ with $i \neq j$. There are $C(6, 2) = 15$ 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

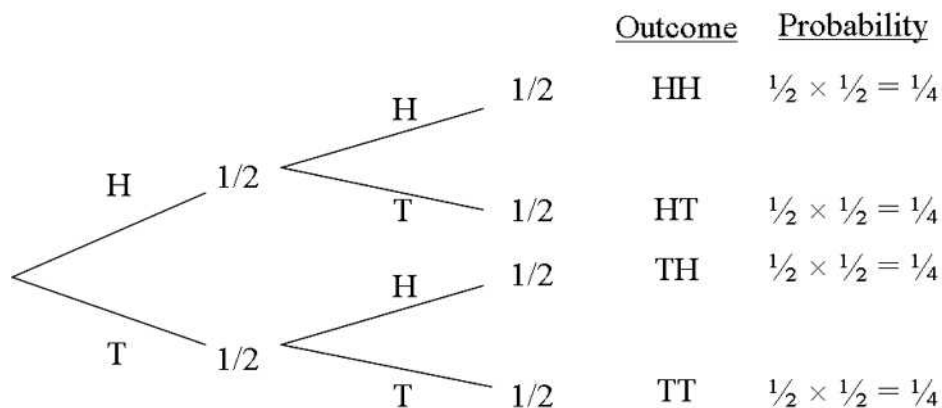


Figure 8.1

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

Suppose that out of 500 computer chips there are 9 defective. Construct the probability tree of the experiment of sampling two of them without replacement.

Solution.

The probability tree is shown in Figure 8.2 ■

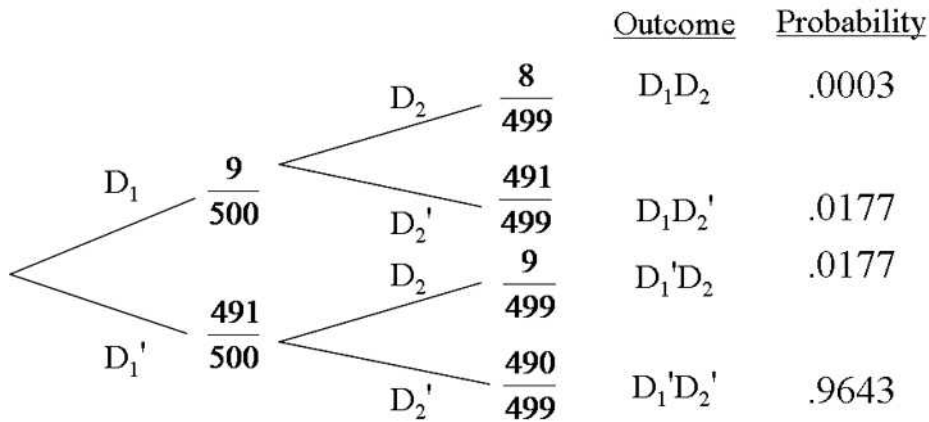


Figure 8.2

Example 8.5

In a state legislature, 35% of the legislators are Democrats, and the other 65% are Republicans. 70% of the Democrats favor raising sales tax, while only 40% of the Republicans favor the increase.

If a legislator is selected at random from this group, what is the probability that he or she favors raising sales tax?

Solution.

Figure 8.3 shows a tree diagram for this problem.

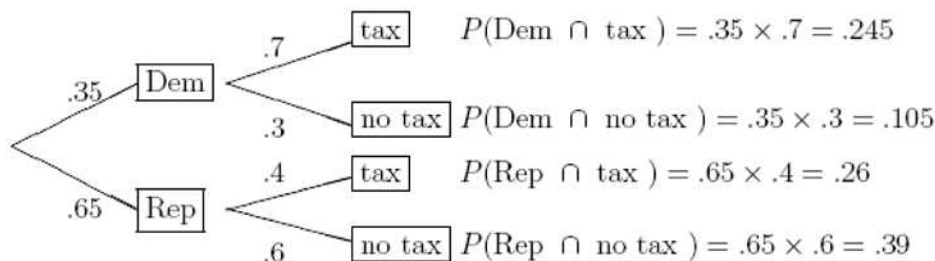


Figure 8.3

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

$$P(\text{tax}) = 0.245 + 0.26 = 0.505 \blacksquare$$

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 the other containing 1,
- (c) submit two batches of 5, one containing 3 fraudulent claims 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) $P(\text{fraud not detected}) = \frac{7}{10} \cdot \frac{6}{9} = \frac{7}{15}$
- (b) $P(\text{fraud not detected}) = \frac{3}{5} \cdot \frac{4}{5} = \frac{12}{25}$
- (c) $P(\text{fraud not detected}) = \frac{3}{5} \cdot 1 = \frac{3}{5}$

Claimant's best strategy is to to split fraudulent claims between two batches of 5 \blacksquare

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

A jar contains three red gumballs and two green gumballs. An experiment consists of drawing gumballs one at a time from the jar, 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?

Problem 8.5

A jar contains three marbles, two black and one red. Two marbles are drawn with replacement. What is the probability that both marbles are black? Assume that the marbles are equally likely to be drawn.

Problem 8.6

A jar contains four marbles—one red, one green, one yellow, and one white. If two marbles are drawn without replacement from the jar, what is the probability of getting a red marble and a white marble?

Problem 8.7

A jar contains 3 white balls and 2 red balls. A ball is drawn at random from the box and not replaced. Then a second ball is drawn from the box. Draw a tree diagram for this experiment and find the probability that the two balls are of different colors.

Problem 8.8

Suppose that a ball is drawn from the box in the previous problem, its color recorded, and then it is put back in the box. Draw a tree diagram for this experiment and find the probability that the two balls are of different colors.

Problem 8.9

Suppose there are 19 balls in an urn. They are identical except in color. 16 of the balls are black and 3 are purple. You are instructed to draw out one ball, note its color, and set it aside. Then you are to draw out another ball and note its color. What is the probability of drawing out a black on the first draw and a purple on the second?

Problem 8.10

A class contains 8 boys and 7 girls. The teacher selects 3 of the children at random and without replacement. Calculate the probability that number of boys selected exceeds the number of girls selected.

Conditional Probability and Independence

In this chapter we introduce the concept of conditional probability. So far, the notation $P(A)$ 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 $P(A|B)$, 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,

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

Thus,

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

provided that $P(B) > 0$.

Example 9.1

Let A denote the event “student is female” and let B denote the event “student is French”. In a class of 100 students suppose 60 are French, and suppose that 10 of the French students are females. Find the probability that if I pick a French student, it will be a female, that is, find $P(A|B)$.

Solution.

Since 10 out of 100 students are both French and female, $P(A \cap B) = \frac{10}{100} = 0.1$. Also, 60 out of the 100 students are French, so $P(B) = \frac{60}{100} = 0.6$. Hence, $P(A|B) = \frac{0.1}{0.6} = \frac{1}{6}$ ■

From

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

we can write

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A). \quad (9.1)$$

Example 9.2

Suppose an individual applying to a college determines that he has an 80% chance of being accepted, and he knows that dormitory housing will only be provided for 60% of all of the accepted students. What is the probability that a student will be accepted and will receive dormitory housing?

Solution.

The probability of the student being accepted and receiving dormitory housing is defined by

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

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

Theorem 9.1

Consider n events A_1, A_2, \dots, A_n . Then

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

Proof.

We have

$$\begin{aligned}
 P(A_1 \cap A_2 \cap \cdots \cap A_n) &= P((A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \cap A_n) \\
 &= P(A_n | A_1 \cap A_2 \cap \cdots \cap A_{n-1}) P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \\
 &= P(A_n | A_1 \cap A_2 \cap \cdots \cap A_{n-1}) P(A_{n-1} | A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \\
 &\quad \times P(A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \\
 &\quad \vdots \\
 &= P(A_n | A_1 \cap A_2 \cap \cdots \cap A_{n-1}) P(A_{n-1} | A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \\
 &\quad \times P(A_{n-2} | A_1 \cap A_2 \cap \cdots \cap A_{n-3}) \cdots \times P(A_2 | A_1) P(A_1)
 \end{aligned}$$

Writing the terms in reverse order we find

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

Example 9.3

Suppose 5 cards are drawn from a pack of 52 cards. Calculate the probability that all cards are the same suit, i.e. a flush.

Solution.

We must find

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

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

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

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

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

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

Theorem 9.2

The function $B \rightarrow P(B | A)$ satisfies the three axioms of ordinary probabilities stated in Section 6.

Proof.

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

$$\begin{aligned} P(\cup_{n=1}^{\infty} B_n | A) &= \frac{P(\cup_{n=1}^{\infty} (B_n \cap A))}{P(A)} \\ &= \frac{\sum_{n=1}^{\infty} P(B_n \cap A)}{P(A)} = \sum_{n=1}^{\infty} P(B_n | A) \blacksquare \end{aligned}$$

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

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

Prior and Posterior Probabilities

The probability $P(A)$ 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 $P(A|B)$ is known as the **posterior probability** of A .

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 $P(A) = \frac{2}{5}$, $P(A \cup B) = \frac{3}{5}$, $P(B|A) = \frac{1}{4}$, $P(C|B) = \frac{1}{3}$, and $P(C|A \cap B) = \frac{1}{2}$. Find $P(A|B \cap C)$.

Problem 9.5

The question, “Do you smoke?” was asked of 100 people. Results are shown in the table.

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

- What is the probability of a randomly selected individual being a male who smokes?
- What is the probability of a randomly selected individual being a male?
- What is the probability of a randomly selected individual smoking?
- What is the probability of a randomly selected male smoking?
- What is the probability that a randomly selected smoker is male?

Problem 9.6

Suppose you have a box containing 22 jelly beans: 10 red, 5 green, and 7 orange. You pick two at random. What is the probability that the first is red and the second is orange?

Problem 9.7

Two fair dice are rolled. What is 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 parts that are either good (90%), slightly defective (2%), or obviously defective (8%). Produced parts get passed through an automatic inspection machine, which is able to detect any part that is obviously defective and discard it. What is the quality of the parts that make it through the inspection machine and get shipped?

Problem 9.9

The probability that a person with certain symptoms has hepatitis is 0.8. A blood 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 hepatitis?

Problem 9.10

If we randomly pick (without replacement) two television tubes in succession from a shipment of 240 television tubes of which 15 are defective, what is the probability that they will both be defective?

Problem 9.11

Find the probabilities of randomly drawing two aces in succession from an ordinary deck of 52 playing cards if we sample

- (a) without replacement
- (b) with replacement

Problem 9.12

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

Problem 9.13

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

Problem 9.14

A computer manufacturer buys cables from three firms. Firm A supplies 50% of all cables and has a 1% defective rate. Firm B supplies 30% of all cables and has a 2% defective rate. Firm C supplies the remaining 20% of cables and has a 5% defective rate.

- (a) What is the probability that a randomly selected cable that the computer manufacturer has purchased is defective (that is what is the overall defective rate of all cables they purchase)?
- (b) Given that a cable is defective, what is the probability it came from Firm A? From Firm B? From Firm C?

Problem 9.15

In a certain town in the United States, 40% of the population are democrats and 60% are republicans. The municipal government has proposed making gun ownership illegal in the town. It is known that 75% of democrats and 30% of republicans support this measure. If a resident of the town is selected at random

- (a) what is the probability that they support the measure?
- (b) If the selected person does support the measure what is the probability the person is a democrat?
- (c) If the person selected does not support the measure, what is the probability that he or she is a democrat?

Problem 9.16

At a certain retailer, purchases of lottery tickets in the final 10 minutes of sale before a draw follow a Poisson distribution with $\lambda = 15$ if the top prize is less than \$10,000,000 and follow a Poisson distribution with $\lambda = 10$ if the top prize is at least \$10,000,000. Lottery records indicate that the top prize is \$10,000,000 or more in 30% of draws.

- (a) Find the probability that exactly 15 tickets are sold in the final 10 minutes before a lottery draw.
- (b) If exactly 15 tickets are sold in the final 10 minutes of sale before the draw, what is the probability that the top prize in that draw is \$10,000,000 or more?

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 $P(A|B)$ we would like to find $P(B|A)$.

Bayes' formula is a simple mathematical formula used for calculating $P(B|A)$ given $P(A|B)$. 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 then we can write

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

Example 10.1

The completion of a construction job may be delayed because of a strike. The probabilities are 0.60 that there will be a strike, 0.85 that the construction job will be completed on time if there is no strike, and 0.35 that the construction will be completed on time if there is a strike. 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 strike. We are given $P(B) = 0.60$, $P(A|B^c) = 0.85$, and $P(A|B) = 0.35$. From Equation (10.1) we find

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

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

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}. \quad (10.2)$$

Example 10.2

A company has two machines A and B for making shoes. It has been observed that machine A produces 10% of the total production of shoes while machine B produces 90% of the total production of shoes. Suppose that 1% of all the shoes produced by A are defective while 5% of all the shoes produced by B are defective. What is the probability that a shoe taken at random from a day's production was made by the machine A , given that it is defective?

Solution.

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

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

Example 10.3

A company that manufactures video cameras produces a basic model (B) and a deluxe model (D). Over the past year, 40% of the cameras sold have been of the basic model. Of those buying the basic model, 30% purchased an extended warranty, whereas 50% of all deluxe purchases do so. If you learn that a randomly selected purchaser has an extended warranty, how likely is it that he or she has a basic model?

Solution.

Let W denote the extended warranty. We are given $P(B) = 0.4$, $P(D) = 0.6$, $P(W|B) = 0.3$, and $P(W|D) = 0.5$. By Bayes' formula we have

$$\begin{aligned} P(B|W) &= \frac{P(B \cap W)}{P(W)} = \frac{P(W|B)P(B)}{P(W|B)P(B) + P(W|D)P(D)} \\ &= \frac{(0.3)(0.4)}{(0.3)(0.4) + (0.5)(0.6)} = 0.286 \blacksquare \end{aligned}$$

Example 10.4

Approximately 1% of women aged 40-50 have breast cancer. A woman with breast cancer has a 90% chance of a positive test from a mammogram, while a woman without breast cancer has a 10% chance of a false positive result. What is the probability a woman has breast cancer given that she just had a positive test?

Solution.

Let B = "the woman has breast cancer" and A = "a positive result." We want to calculate $P(B|A)$ given by

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

But

$$P(B \cap A) = P(A|B)P(B) = (0.9)(0.01) = 0.009$$

and

$$P(B^c \cap A) = P(A|B^c)P(B^c) = (0.1)(0.99) = 0.099.$$

Thus,

$$P(B|A) = \frac{0.009}{0.009 + 0.099} = \frac{9}{108} \blacksquare$$

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, \dots, H_n with $P(H_i) > 0$ for each i . Then for any event A and $1 \leq i \leq n$ we have

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

where

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + \dots + P(H_n)P(A|H_n).$$

Proof.

First note that

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

Hence,

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

Example 10.5

Suppose a voter poll is taken in three states. In state A, 50% of voters support the liberal candidate, in state B, 60% of the voters support the liberal candidate, and in state C, 35% of the voters support the liberal candidate. Of the total population of the three states, 40% live in state A, 25% live in state B, and 35% live in state C. Given that a voter supports the liberal candidate, what is the probability that he/she lives in state B?

Solution.

Let L_I denote the event that a voter lives in state I, where I = A, B, or C. Let S denote the event that a voter supports liberal candidate. We want to find $P(L_B|S)$. By Bayes' formula we have

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

Example 10.6

The members of a consulting firm rent cars from three rental companies: 60% from company 1, 30% from company 2, and 10% from company 3. Past statistics show that 9% of the cars from company 1, 20% of the cars from company 2, and 6% of the cars from company 3 need a tune-up. If a rental car delivered to the firm needs a tune-up, what is the probability that it came from company 2?

Solution.

Define the events

H_1 = car comes from company 1

H_2 = car comes from company 2

H_3 = car comes from company 3

T = car needs tune up

Then

$$\begin{aligned} P(H_1) &= 0.6 & P(H_2) &= 0.3 & P(H_3) &= 0.1 \\ P(T|H_1) &= 0.09 & P(T|H_2) &= 0.2 & P(T|H_3) &= 0.06 \end{aligned}$$

From Bayes' theorem we have

$$\begin{aligned} P(H_2|T) &= \frac{P(T|H_2)P(H_2)}{P(T|H_1)P(H_1) + P(T|H_2)P(H_2) + P(T|H_3)P(H_3)} \\ &= \frac{0.2 \times 0.3}{0.09 \times 0.6 + 0.2 \times 0.3 + 0.06 \times 0.1} = 0.5 \blacksquare \end{aligned}$$

Example 10.7

The probability that a person with certain symptoms has hepatitis is 0.8. A blood 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 hepatitis ?

Solution.

Let H denote the event that a person (with these symptoms) has Hepatitis, H^c the event that they do not, and Y the event that the test results are positive. We have $P(H) = 0.8$, $P(H^c) = 0.2$, $P(Y|H) = 0.9$, $P(Y|H^c) = 0.05$. We want $P(H|Y)$, and this is found by Bayes' formula

$$P(H|Y) = \frac{P(Y|H)P(H)}{P(Y|H)P(H) + P(Y|H^c)P(H^c)} = \frac{(0.9)(0.8)}{(0.9)(0.8) + (0.05)(0.2)} = \frac{72}{73} \blacksquare$$

Example 10.8

A factory produces its products with three machines. Machine I, II, and III produces 50%, 30%, and 20% of the products, but 4%, 2%, and 4% of their products are defective, respectively.

- (a) What is the probability that a randomly selected product is defective?
 (b) If a randomly selected product was found to be defective, what is the probability that this product was produced by machine I?

Solution.

Let I, II, and III denote the events that the selected product is produced by machine I, II, and III, respectively. Let D be the event that the selected product is defective. Then, $P(I) = 0.5$, $P(II) = 0.3$, $P(III) = 0.2$, $P(D|I) = 0.04$, $P(D|II) = 0.02$, $P(D|III) = 0.04$. So, by the total probability rule, we have

$$\begin{aligned} P(D) &= P(D|I)P(I) + P(D|II)P(II) + P(D|III)P(III) \\ &= (0.04)(0.50) + (0.02)(0.30) + (0.04)(0.20) = 0.034 \end{aligned}$$

- (b) By Bayes' theorem, we find

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

Example 10.9

A building has 60 occupants consisting of 15 women and 45 men. The men have probability $1/2$ of being colorblind and the women have probability $1/3$ of being colorblind.

- (a) Suppose you choose uniformly at random a person from the 60 in the building. What is the probability that the person will be colorblind?
 (b) Determine the conditional probability that you chose a woman given that the occupant you chose is colorblind.

Solution.

Let

$W = \{\text{the one selected is a woman}\}$

$M = \{\text{the one selected is a man}\}$

$B = \{\text{the one selected is color – blind}\}$

(a) We are given the following information: $P(W) = \frac{15}{60} = \frac{1}{4}$, $P(M) = \frac{3}{4}$, $P(B|W) = \frac{1}{3}$, and $P(B|M) = \frac{1}{2}$. By the total law of probability we have

$$P(B) = P(B|W)P(W) + P(B|M)P(M) = \frac{11}{24}.$$

(b) Using Bayes' theorem we find

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

Problems

Problem 10.1

An insurance company believes that people can be divided into two classes: those who are accident prone and those who are not. Their statistics show that an accident prone person will have an accident at some time within 1-year period with probability 0.4, whereas this probability decreases to 0.2 for a non-accident prone person.

- (a) If we assume that 30% of the population is accident prone, 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 accident prone?

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

In a certain state, 25% of all cars emit excessive amounts of pollutants. If the probability is 0.99 that a car emitting excessive amounts of pollutants will fail the state's vehicular emission test, and the probability is 0.17 that a car not emitting excessive amounts of pollutants will nevertheless fail the test, what is the probability that a car that fails the test actually emits excessive amounts of pollutants?

Problem 10.11

I haven't decided what to do over spring break. There is a 50% chance that

I'll go skiing, a 30% chance that I'll go hiking, and a 20% chance that I'll stay home and play soccer. The (conditional) probability of my getting injured is 30% if I go skiing, 10% if I go hiking, and 20% if I play soccer.

- (a) What is the probability that I will get injured over spring break?
- (b) If I come back from vacation with an injury, what is the probability that I got it skiing?

Problem 10.12

The final exam for a certain math class is graded pass/fail. A randomly chosen student from a probability class has a 40% chance of knowing the material well. If he knows the material well, he has an 80% chance of passing the final. If he doesn't know the material well, he has a 40% chance of passing the final anyway.

- (a) What is the probability of a randomly chosen student passing the final exam?
- (b) If a student passes, what is the probability that he knows the material?

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

Suppose 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 Math 408 as those who had a lower grade in calculus. If a student who received an A in Math 408 is chosen at random, what is the probability that he/she also received an A in calculus?

(Assume all students in Math 408 have taken calculus.)

Problem 10.15

A designer buys fabric from 2 suppliers. Supplier A provides 60% of the fabric, supplier B provides the remaining 40%. Supplier A has a defect rate of 5%. Supplier B has a defect rate of 8%. Given a batch of fabric is defective, what is the probability it came from Supplier A?

Problem 10.16

$1/10$ of men and $1/7$ of women are color-blind. A person is randomly selected. Assume males and females to be in equal numbers.

- (a) What is the probability that the selected person is color-blind?
- (b) If the selected person is color-blind, what is the probability that the person is male?

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

$$P(A|B) = P(A).$$

We next introduce the two most basic theorems regarding independence.

Theorem 11.1

A and B are independent events if and only if $P(A \cap B) = P(A)P(B)$.

Proof.

A and B are independent if and only if $P(A|B) = P(A)$ and this is equivalent to

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

Example 11.1

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

Solution.

We have

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

Example 11.2

An oil exploration company currently has two active projects, one in Asia and the other in Europe. Let A be the event that the Asian project is successful and B the event that the European project is successful. Suppose that A and B are independent events with $P(A) = 0.4$ and $P(B) = 0.7$. What is the probability that at least one of the two projects will be successful?

Solution.

The probability that at least one of the two projects is successful is $P(A \cup B)$. Thus,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)P(B) = 0.82 \blacksquare$$

Example 11.3

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

Solution.

First, note that

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

Next,

$$P(B|A \cup B) = \frac{P(B)}{P(A \cup B)} = \frac{P(B)}{P(A) + P(B) - P(A \cap B)} = \frac{P(B)}{P(A) + P(B) - P(A)P(B)}.$$

Thus,

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

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

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, $P(A) = P(A \cap B) + P(A \cap B^c)$. It follows that

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

Example 11.4

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

Solution.

Using De Morgan's formula we have

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

Example 11.5

A probability for a woman to have twins is 5%. A probability to win the lottery is 1:1,000,000. Joan has twins. What is the probability for her to win the lottery?

Solution.

Since the two events are independent, the probability for Joan to win the lottery is 1:1,000,000 ■

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 11.6

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.

Since $P(A) = P(B) = \frac{13}{52} = \frac{1}{4}$ and

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

then 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, \dots, A_n are said to be **mutually independent** or simply **independent** if for any $1 \leq i_1 < i_2 < \dots < i_k \leq n$ we have

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k})$$

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

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

Example 11.7

Consider the experiment of tossing a coin n times. Let $A_i =$ “the i^{th} coin shows Heads”. Show that A_1, A_2, \dots, A_n are independent.

Solution.

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

Example 11.8

The probability that a lab specimen contains high levels of contamination is 0.05. Four samples are checked, and the samples are independent.

- (a) What is the probability that none contains high levels of contamination?
- (b) What is the probability that exactly one contains high levels of contamination?
- (c) What is the probability that at least one contains high levels of contamination?

Solution.

Let H_i denote the event that the i^{th} sample contains high levels of contamination for $i = 1, 2, 3, 4$.

The event that none contains high levels of contamination 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

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

(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, $P(A_i) = (0.95)^3(0.05) = 0.0429, i = 1, 2, 3, 4$. Therefore, the answer is $4(0.0429) = 0.1716$.

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

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

Example 11.9

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, \dots, A_n are said to be **pairwise independent** if and only if $P(A_i \cap A_j) = P(A_i)P(A_j)$ for any $i \neq j$ where $1 \leq i, j \leq n$.

Pairwise independence does not imply independence as the following example shows.

Example 11.10

The four sides of a tetrahedron (regular three sided pyramid with 4 sides consisting of isosceles triangles) are denoted by 2, 3, 5 and 30, respectively. If the tetrahedron is rolled the number on the base is the outcome of interest. Consider the three events A = "the number on the base is even", B = "the number is divisible by 3", and C = "the number is divisible by 5". Show that these events are pairwise independent, but not independent.

Solution.

Note that $A = \{2, 30\}$, $B = \{3, 30\}$, and $C = \{5, 40\}$. Then

$$P(A \cap B) = P(\{30\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A)P(B)$$

$$P(A \cap C) = P(\{30\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A)P(C)$$

$$P(B \cap C) = P(\{30\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(B)P(C)$$

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

$$P(A \cap B \cap C) = P(\{30\}) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(A)P(B)P(C)$$

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

Problems

Problem 11.1

Determine whether the events are independent or dependent.

- (a) Selecting a marble and then choosing a second marble without replacing the first marble.
- (b) Rolling a number cube and spinning a spinner.

Problem 11.2

David and Adrian have a coupon for a pizza with one topping. The choices of toppings are pepperoni, hamburger, sausage, onions, bell peppers, olives, and anchovies. If they choose at random, what is the probability that they both choose hamburger 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

You and two friends go to a restaurant and order a sandwich. The menu has 10 types of sandwiches and each of you is equally likely to order any type. What is the probability that each of you orders a different type?

Problem 11.5 ‡

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. The probability that both balls are the same color is 0.44 . Calculate the number of blue balls in the second urn.

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 consists only of 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 $P(A) = 0.2$ and $P(B) = 0.3$. 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 $P(C)$ and $P(D)$.

Problem 11.10

Suppose A, B , and C are mutually independent events with probabilities $P(A) = 0.5$, $P(B) = 0.8$, and $P(C) = 0.3$. 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 $P(A) = 0.5$, $P(B) = 0.8$, and $P(C) = 0.3$. 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 $P(A)$, $P(B)$ and $P(C)$.
- (c) Compute $P(B|A)$.
- (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 and/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" 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 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,

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

and

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

Example 12.1

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

Solution.

We have

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

Converting Probability to Odds

Given $P(E)$, we want to find the odds in favor of E and the odds against E . The odds in favor of E are

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

and the odds against E are

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

Example 12.2

For each of the following, find the odds in favor of the event's occurring:

- Rolling a number less than 5 on a die.
- Tossing heads on a fair coin.
- 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 $P(H) = \frac{1}{2}$ and $P(T) = \frac{1}{2}$ then 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 $P(\text{ace}) = \frac{4}{52}$ and $P(\text{not an ace}) = \frac{48}{52}$ 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 $P(E) = \frac{5}{6}$ is just a ratio. The exact number of favorable outcomes and the exact total of all outcomes are not necessarily known.

Now consider a hypothesis H that is true with probability $P(H)$ and suppose that new evidence E is introduced. Then the conditional probabilities, given the evidence E , that H is true and that H is not true are given by

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)} \text{ and } P(H^c|E) = \frac{P(E|H^c)P(H^c)}{P(E)}.$$

Therefore, the new odds after the evidence E has been introduced is

$$\frac{P(H|E)}{P(H^c|E)} = \frac{P(H)}{P(H^c)} \frac{P(E|H)}{P(E|H^c)}.$$

That is, the new value of the odds of H is its old value multiplied by the ratio of the conditional probability of the new evidence given that H is true to the conditional probability given that H is not true.

Example 12.3

The probability that a coin C_1 comes up heads is $\frac{1}{4}$ and that of a coin C_2 is $\frac{3}{4}$. Suppose that one of these coins is randomly chosen and is flipped twice. If both flips land heads, what are the odds that coin C_2 was the one flipped?

Solution.

Let H be the event that coin C_2 was the one flipped and E the event a coin flipped twice lands two heads. Since $P(H) = P(H^c)$

$$\begin{aligned} \frac{P(H|E)}{P(H^c|E)} &= \frac{P(H)}{P(H^c)} \frac{P(E|H)}{P(E|H^c)} = \frac{P(E|H)}{P(E|H^c)} \\ &= \frac{\frac{9}{16}}{\frac{1}{16}} = 9. \end{aligned}$$

Hence, the odds are 9:1 that coin C_2 was the one flipped ■

Problems

Problem 12.1

If the probability of a boy's 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 Deborah'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 rain for the day is 60%, what are the odds against its raining?

Problem 12.5

On a tote board at a race track, the odds for Spiderman 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 Spiderman'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 $P(E) = \frac{3}{4}$.

Problem 12.8

Find $P(E)$ in each case.

- (a) The odds in favor of E are 3:4
- (b) The odds against E are 7:3

Problem 12.9

A financial analyst states that based on his economic information, that the odds of a recession in the U.S. in the next two years is 2:1. State his beliefs in terms of the probability of a recession in the U.S. in the next two years.

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** random variable is usually the result of a count and therefore the range consists of whole integers. A **continuous** random variable is usually the result of a measurement. As a result the range is any subset of the set of all real numbers. 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 or continuous.

- (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.

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 ■

Example 13.2

Toss a coin 3 times: $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.
 Let $X = \#$ of Heads in 3 tosses. Find the range of X .

Solution.

We have

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

Thus, the range of X consists of $\{0, 1, 2, 3\}$ ■

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 $\{HHT, HTH, TTH\}$. 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, $P(X = 2) = \frac{3}{8}$.

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 2 rolls. Complete the following table

x	1	2	3	4
P(X=x)				

Solution.

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

Example 13.4

Five male and five female are ranked according to their scores on an examination. Assume that no two scores are alike and all $10!$ possible rankings are equally likely. Let X denote the highest ranking achieved by a female (for instance, $X = 1$ if the top-ranked person is female). Find $P(X = i), i = 1, 2, \dots, 10$.

Solution.

Since 6 is the lowest possible rank attainable by the highest-scoring female, we must have $P(X = 7) = P(X = 8) = P(X = 9) = P(X = 10) = 0$.

For $X = 1$ (female is highest-ranking scorer), we have 5 possible choices out of 10 for the top spot that satisfy this requirement; hence $P(X = 1) = \frac{1}{2}$.

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

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

For $X = 3$ (female is 3rd-highest scorer), acceptable configurations yield $(5)(4) = 20$ possible choices for the top 2 males, 5 possible choices for the female 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,

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

Similarly, we have

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

Problems

Problem 13.1

Determine whether the random variable is discrete or continuous.

- (a) Time between oil changes on a car
- (b) Number of heart beats per minute
- (c) The number of calls at a switchboard in a day
- (d) Time it takes to finish a 60-minute exam

Problem 13.2

Two socks are selected at random and removed in succession from a drawer containing five brown socks and three green socks. 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 brown socks 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 $P(X = 6)$.

Problem 13.4

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

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

Find $P(X < 50)$.

Problem 13.5

Toss a fair coin until the coin lands on heads. Let the random variable X denote the number of coin flips until the first head appears. Find $P(X = n)$ where n is a positive integer.

Problem 13.6

Choose a baby name at random from $S = \{ \text{Josh, John, Thomas, Peter} \}$. Let $X(\omega) = \text{first letter in name}$. Find $P(X = J)$.

Problem 13.7 ‡

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

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

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
$P(X=x)$	0.02	0.41	0.21	0.08	0.28

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

Problem 13.9

A basketball player shoots three times, independently. The success probability of his first, second, and third shots are 0.7, 0.5, and 0.4 respectively. Let X denote the number of successful shots among these three.

- (a) Define the function X from the sample space S into \mathbb{R} .
- (b) $\{X = 0\}$ corresponds to what subset in S ? What is the probability that he misses all three shots; i.e $P(X = 0)$?
- (c) $\{X = 1\}$ corresponds to what subset in S ? What is the probability that he succeeds exactly once among these three shots; i.e $P(X = 1)$?
- (d) $\{X = 2\}$ corresponds to what subset in S ? What is the probability that he succeeds exactly twice among these three shots; i.e $P(X = 2)$?
- (e) $\{X = 3\}$ corresponds to what subset in S ? What is the probability that he makes all three shots; i.e $P(X = 3)$?

Problem 13.10

Four distinct numbers are randomly distributed to players numbered 1 through 4. Whenever two players compare their numbers, the one with the higher one is declared the winner. Initially, players 1 and 2 compare their numbers; the winner then compares with player 3, and so on. Let X denote the number of times player 1 is a winner. Find $P(X = i)$, $i = 0, 1, 2, 3$.

Problem 13.11

For a certain discrete random variable on the non-negative integers, the probability function satisfies the relationships

$$P(0) = P(1), \quad P(k+1) = \frac{1}{k}P(k), \quad k = 1, 2, 3, \dots$$

Find $P(0)$.

14 Probability Mass Function and Cumulative Distribution Function

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

$$p(x) = P(X = x).$$

That is, a probability mass function (pmf) 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 ■

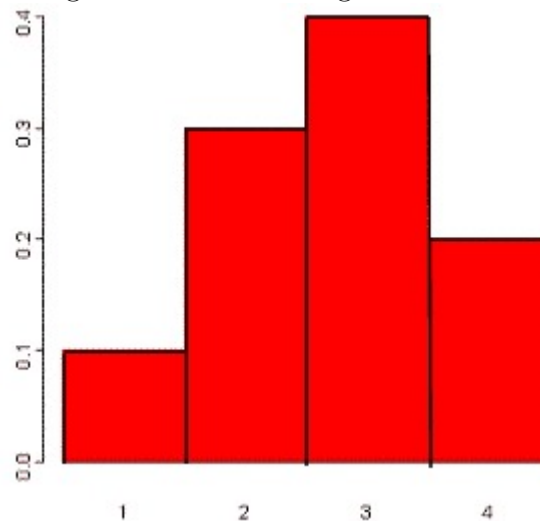


Figure 14.1

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

$$p(x) = \frac{\binom{5}{x} \binom{5}{4-x}}{\binom{10}{4}}.$$

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 four-sided die twice. Let $X(\omega_1, \omega_2) = \max\{\omega_1, \omega_2\}$. Find the equation of $p(x)$.

Solution.

The pmf of X is

$$\begin{aligned} p(x) &= \begin{cases} \frac{2x-1}{16} & \text{if } x = 1, 2, 3, 4 \\ 0 & \text{otherwise} \end{cases} \\ &= \frac{2x-1}{16} I_{\{1,2,3,4\}}(x) \end{aligned}$$

where

$$I_{\{1,2,3,4\}}(x) = \begin{cases} 1 & \text{if } x \in \{1, 2, 3, 4\} \\ 0 & \text{otherwise} \blacksquare \end{cases}$$

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

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

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

$$\begin{aligned} p(x) &\geq 0, & x \in \Omega \\ p(x) &= 0, & x \notin \Omega \end{aligned}$$

Moreover,

$$\sum_{x \in \Omega} p(x) = 1.$$

All random variables (discrete and continuous) have a **distribution function** or **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,

$$F(a) = P(X \leq a) = \sum_{x \leq a} p(x).$$

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 ■

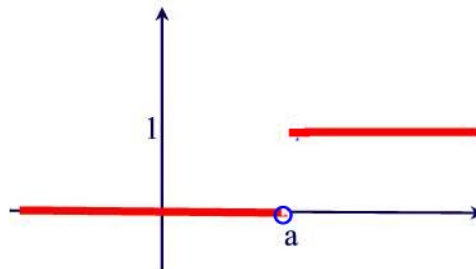


Figure 14.2

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.

Example 14.5

Consider the following probability 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 ■

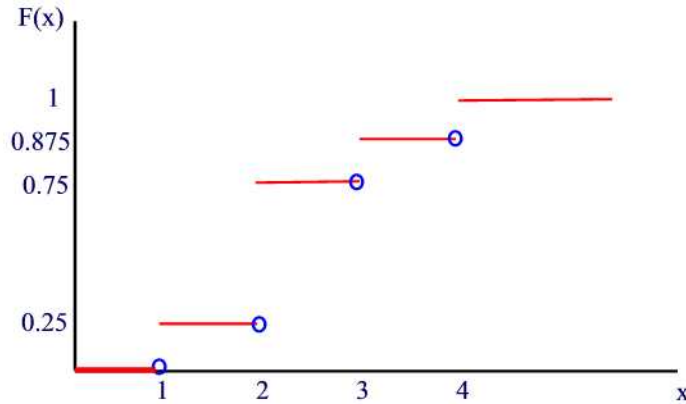


Figure 14.3

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 $x_1 < x_2 < \dots < x_n$ then $p(x_1) = F(x_1)$ and

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

Proof.

Because $F(x) = 0$ for $x < x_1$ then $F(x_1) = P(X \leq x_1) = P(X < x_1) + P(X = x_1) = P(X = x_1) = p(x_1)$. Also, by Proposition 21.7 of Section 21, for $2 \leq i \leq n$ we can write $p(x_i) = P(x_{i-1} < X \leq x_i) = F(x_i) - F(x_{i-1})$ ■

Example 14.6

If the 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 ■

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

Take a carton of six machine parts, with one defective and five nondefective. Randomly draw parts for inspection, but do not replace them after each draw, until a defective part fails inspection. Let X denote the number of parts drawn until a nondefective part is found. Find the probability mass function of X and draw its histogram.

Problem 14.5

Roll two dice, and let the random variable X denote 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

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

Find a formula for $F(n)$.

Problem 14.7

A box contains 100 computer chips, of which exactly 5 are good and the remaining 95 are bad.

(a) Suppose first you take out chips one at a time (without replacement) and test each chip you have taken out until you have found a good one. Let X be the number of chips you have to take out in order to find one that is good. (Thus, X can be as small as 1 and as large as 96.) Find the probability distribution of X .

(b) Suppose now that you take out exactly 10 chips and then test each of these 10 chips. Let Y denote the number of good chips among the 10 you have 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 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 your gain/lost. Find the probability mass function of X .

Problem 14.10

An experiment consists of randomly tossing a biased coin 3 times. The probability of heads on any particular toss is known to be $\frac{1}{3}$. Let X denote the number of heads.

(a) Find the probability mass function.

(b) Plot the probability mass distribution function for X .

Problem 14.11

If the distribution function of X is given by

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

find the probability distribution of X .

Problem 14.12

A state lottery game draws 3 numbers at random (without replacement) from a set of 15 numbers. Give the probability distribution for the number of “winning digits” you will have when you purchase one ticket.

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

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

Thus, $P(L) = \frac{11}{18} = 61\%$. 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, \dots, x_k , and let $p(x)$ be the corresponding probability mass function. Then the **expected value** of X is

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

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

$$p_i = P(X = x_i) = p(x_i), i = 1, 2, \dots, k.$$

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

$$n_1x_1 + n_2x_2 + \dots + n_kx_k$$

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 $P(X = x_i) = \lim_{n \rightarrow \infty} \frac{n_i}{n}$. 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

An American roulette wheel has 38 compartments around its rim. Two of these are colored green and are numbered 0 and 00. The remaining compartments are numbered 1 through 36 and are alternately colored black and red. When the wheel is spun in one direction, a small ivory ball is rolled in

the opposite direction around the rim. When the wheel and the ball slow down, the ball eventually falls in any one of the compartments with equal likelihood if the wheel is fair. One way to play is to bet on whether the ball will fall in a red slot or a black slot. If you bet on red for example, you win the amount of the bet if the ball lands in a red slot; otherwise you lose. What is the expected win if you consistently bet \$5 on red?

Solution.

Let X denote the winnings of the game. The probability of winning is $\frac{18}{38}$ and that of losing is $\frac{20}{38}$. Your expected win is

$$E(X) = \frac{18}{38} \times 5 - \frac{20}{38} \times 5 \approx -0.26$$

On average you should expect to lose 26 cents per play ■

Example 15.3

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 $p(1) = P(A)$ and $p(0) = p(A^c)$ then

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

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

Example 15.4

A man determines that the probability of living 5 more years is 0.85. His insurance policy pays \$1,000 if he dies within the next 5 years. Let X be the random variable that represents the amount the insurance company pays out in the next 5 years.

- What is the probability distribution of X ?
- What is the most he should be willing to pay for the policy?

Solution.

(a) $P(X = 1000) = 0.15$ and $P(X = 0) = 0.85$.

(b) $E(X) = 1000 \times 0.15 + 0 \times 0.85 = 150$. Thus, his expected payout is \$150, so he should not be willing to pay more than \$150 for the policy ■

Remark 15.2

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 $p(x)$ 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 centre 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.

Problems

Problem 15.1

Compute the expected value of the sum of two faces when rolling two dice.

Problem 15.2

A game consists of rolling two dice. You win the amounts shown for rolling the score shown.

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 the game.

Problem 15.3

You play a game in which two dice are rolled. If a sum of 7 appears, you win \$10; otherwise, you lose \$2.00. If you intend to play this game for a long time, should you expect to make money, lose money, or come out about even? Explain.

Problem 15.4

Suppose it costs \$8 to roll a pair of dice. You get paid the sum of the numbers in dollars that appear on the dice. What is the expected value of this game?

Problem 15.5

Consider the spinner in Figure 15.1, with the payoff in each sector of the circle. Should the owner of this spinner expect to make money over an extended period of time if the charge is \$2.00 per spin?

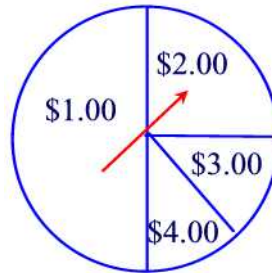


Figure 15.1

Problem 15.6

An insurance company will insure your dorm room against theft for a semester. Suppose the value of your possessions is \$800. The probability of your being robbed of \$400 worth of goods during a semester is $\frac{1}{100}$, and the probability of your being robbed of \$800 worth of goods is $\frac{1}{400}$. Assume that these are the only possible kinds of robberies. How much should the insurance company charge people like you to cover the money they pay out and to make an additional \$20 profit per person on the average?

Problem 15.7

Consider a lottery game in which 7 out of 10 people lose, 1 out of 10 wins \$50, and 2 out of 10 wins \$35. If you played 10 times, about how much would you expect to win? Assume that a game costs a dollar.

Problem 15.8

You pick 3 different numbers between 1 and 12. If you pick all the numbers correctly you win \$100. What are your expected earnings if it costs \$1 to play?

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

A professor has made 30 exams of which 8 are “hard”, 12 are “reasonable”, and 10 are “easy”. The exams are mixed up and the professor selects randomly four of them without replacement to give to four sections of the course she is teaching. Let X be the number of sections that receive a hard test.

- What is the probability that no sections receive a hard test?
- What is the probability that exactly one section receives a hard test?
- Compute $E(X)$.

Problem 15.11

Consider a random variable X whose cumulative distribution function is given by

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

We are also told that $P(X > 3) = 0.1$.

- What is q ?
- Compute $P(X^2 > 2)$.
- What is $p(0)$? What is $p(1)$? What is $p(P(X \leq 0))$? (Here, $p(x)$ denotes the probability mass function (pmf) for X)
- Find the formula of the function $p(x)$.
- Compute $E(X)$.

Problem 15.12

Used cars can be classified in two categories: peaches and lemons. Sellers know which type the car they are selling is. Buyers do not. Suppose that buyers are aware that the probability a car is a peach is 0.4. Sellers value peaches at \$2500, and lemons at \$1500. Buyers value peaches at \$3000, and lemons at \$2000.

- Obtain the expected value of a used car to a buyer who has no extra information.
- Assuming that buyers will not pay more than their expected value for a used car, will sellers ever sell peaches?

Problem 15.13

In a batch of 10 computer parts it is known that there are three defective parts. Four of the parts are selected at random to be tested. Define the random variable X to be the number of working (non-defective) computer parts selected.

- Derive the probability mass function of X .
- What is the cumulative distribution function of X ?
- Find the expected value and variance of X .

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 $P(X = -1) = 0.2$, $P(X = 0) = 0.5$, and $P(X = 1) = 0.3$. Compute $E(X^2)$.

Solution.

Let $Y = X^2$. Then the range of Y is $\{0, 1\}$. Also, $P(Y = 0) = P(X = 0) = 0.5$ and $P(Y = 1) = P(X = -1) + P(X = 1) = 0.2 + 0.3 = 0.5$. 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 $p(x)$, 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 $\{s \in S : g(X)(s) = y\} = \cup_{x \in A_y} \{s \in S : X(s) = x\}$. 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,

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

Now, from the definition of the expected value we have

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

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

Example 16.2

If X is the number of points rolled with a balanced 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 the mean and variance of Y .

Solution.

By the properties of expectation,

$$\begin{aligned} E(Y) &= E(20 - 2X) = 20 - 2E(X) = 20 - 12 = 8 \\ E(Y^2) &= E(400 - 80X + 4X^2) = 400 - 80E(X) + 4E(X^2) = 100 \\ \text{Var}(Y) &= E(Y^2) - (E(Y))^2 = 100 - 64 = 36 \blacksquare \end{aligned}$$

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 **nth moment** of X . 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

$$\begin{aligned} E(X^2) &= \sum_{x \in D} x^2 p(x) \\ &= \sum_{x \in D} (x(x - 1) + x) p(x) \\ &= \sum_{x \in D} x(x - 1) p(x) + \sum_{x \in D} x p(x) = E(X(X - 1)) + E(X) \blacksquare \end{aligned}$$

Remark 16.1

In our definition of expectation the set D can be infinite. It is possible to have a random variable with undefined expectation (See Problem 16.11)

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

A die is loaded in such a way that the probability of any particular face showing is directly proportional to the number on that face. Let X denote the number showing for a single toss of this die.

- (a) What is the probability function $p(x)$ for X ?
- (b) What is the probability that an even number occurs on a single roll of this die?
- (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 $N = 1, \dots, 5$ 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 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.

- Find the probability mass function of X .
- Compute $E(2^X)$.

Problem 16.10

A movie theater sell three kinds of tickets - children (for \$3), adult (for \$8),

and seniors (for \$5). The number of children tickets has $E[C] = 45$. The number of adult tickets has $E[A] = 137$. Finally, the number of senior tickets has $E[S] = 34$. You may assume the number of tickets in each group is independent.

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 $E(P)$.

Problem 16.11

If the probability distribution of X is given by

$$p(x) = \left(\frac{1}{2}\right)^x, \quad x = 1, 2, 3, \dots$$

show that $E(2^X)$ does not exist.

Problem 16.12

An industrial salesperson has a salary structure as follows:

$$S = 100 + 50Y - 10Y^2$$

where Y is the number of items sold. Assuming sales are independent from one attempt to another, each with probability of success of $p = 0.20$. Give the expected salaries on days where the salesperson attempts $n = 1, 2$, and 3.

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 $SD(X)$ denotes the standard deviation then the variance is given by the formula

$$\text{Var}(X) = SD(X)^2 = E[(X - E(X))^2]$$

The variance of a random variable is typically calculated using the following formula

$$\begin{aligned} \text{Var}(X) &= E[(X - E(X))^2] \\ &= E[X^2 - 2XE(X) + (E(X))^2] \\ &= E(X^2) - 2E(X)E(X) + (E(X))^2 \\ &= E(X^2) - (E(X))^2 \end{aligned}$$

where we have used the result of Problem 16.4.

Example 17.1

We toss a fair coin and let $X = 1$ if we get heads, $X = -1$ if we get tails. Find the variance of X .

Solution.

Since $E(X) = 1 \times \frac{1}{2} - 1 \times \frac{1}{2} = 0$ and $E(X^2) = 1^2 \times \frac{1}{2} + (-1)^2 \times \frac{1}{2} = 1$ then $\text{Var}(X) = 1$ ■

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

$$\text{Var}(aX + b) = a^2\text{Var}(X)$$

Proof.

Since $E(aX + b) = aE(X) + b$ then

$$\begin{aligned}\text{Var}(aX + b) &= E[(aX + b - E(aX + b))^2] \\ &= E[a^2(X - E(X))^2] \\ &= a^2 E((X - E(X))^2) \\ &= a^2 \text{Var}(X) \blacksquare\end{aligned}$$

Remark 17.1

Note that the units of $\text{Var}(X)$ is the square of the units of X . This motivates the definition of the standard deviation $\sigma = \sqrt{\text{Var}(X)}$ which is measured in the same units as X .

Example 17.2

This year, Toronto Maple Leafs tickets cost an average of \$80 with a variance of \$105. Toronto city council wants to charge a 3% tax on all tickets(i.e., all tickets will be 3% more expensive. If this happens, what would be the variance of the cost of Toronto Maple Leafs 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

Roll one die and let X be the resulting number. 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

$$SD(X) = \sqrt{\frac{35}{12}} \approx 1.7078 \blacksquare$$

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 $g(c) \geq 0$ and it is minimum when $c = E(X)$ ■

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

A recent study indicates that the annual cost of maintaining and repairing a car in a town in Ontario averages 200 with a variance of 260.

If a tax of 20% is introduced on all items associated with the maintenance and repair of cars (i.e., everything is made 20% more expensive), what will be the variance of the annual cost of maintaining and repairing a car?

Problem 17.3

A discrete random variable, X , has probability mass function

$$p(x) = c(x - 3)^2, \quad x = -2, -1, 0, 1, 2.$$

- Find the value of the constant c .
- Find the mean and variance of X .

Problem 17.4

In a batch of 10 computer parts it is known that there are three defective parts. Four of the parts are selected at random to be tested. Define the random variable X to be the number of working (non-defective) computer parts selected.

- Derive the probability mass function of X .
- Find the expected value and 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 $\text{Var}(X)$.

Problem 17.6

Suppose X is a random variable with $E(X) = 4$ and $\text{Var}(X) = 9$. Let $Y = 4X + 5$. Compute $E(Y)$ and $\text{Var}(Y)$.

Problem 17.7

Many students come to my office during the first week of class with issues related to registration. The students either ask about STA200 or STA291 (never both), and they can either want to add the class or switch sections (again, not both). Suppose that 60% of the students ask about STA200. Of those who ask about STA200, 70% want to add the class. Of those who ask about STA291, 40% want to add the class.

- (a) Construct a probability table.
- (b) Of those who want to switch sections, what proportion are asking about STA291?
- (c) Suppose it takes 15 minutes to deal with “STA200 and adding”, 12 minutes to deal with “STA200 and switching”, 8 minutes to deal with “STA291 and adding”, and 7 minutes to deal with “STA291 and switching”. Let X be the amount of time spend with a particular student. Find the probability mass function of X ?
- d) What are the expectation and variance of X ?

Problem 17.8

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 loose \$ 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 $\text{Var}(X)$.

Problem 17.9

Let X be a discrete random variable with probability mass function given in tabular form.

x	-4	1	4
P(x)	$\frac{3}{10}$	$\frac{4}{10}$	$\frac{3}{10}$

Find the variance and the standard deviation of X .

Problem 17.10

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 $Var(X)$.

18 Binomial and Multinomial Random Variables

Binomial experiments are problems that consist of a fixed number of trials n , with each trial having exactly two possible 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

$$p + q = 1.$$

Also, we assume that the trials are **independent**, that is what happens to one trial does not affect the probability of a success in any other trial.

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 in the binomial experiment.

Example 18.1

Consider the experiment of tossing a fair coin.

- (a) What makes a trial?
- (b) What is a success? a failure?

Solution.

- (a) A trial consists of tossing the coin.
- (b) A success could be a Head and a failure is a Tail ■

Let X represent the number of successes that occur in n trials. Then X is said to be a **binomial random variable** with parameters (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

Privacy is a concern for many users of the Internet. One survey showed that 79% of Internet users are somewhat concerned about the confidentiality of their e-mail. Based on this information, we would like to find the probability that for a random sample of 12 Internet users, 7 are concerned about the privacy of their e-mail. What are the values of n, p, q, r ?

Solutions.

This is a binomial experiment with 12 trials. If we assign success to an Internet user being concerned about the privacy of e-mail, the probability of success is 79%. We are interested in the probability of 7 successes. We have $n = 12, p = 0.79, q = 1 - 0.79 = 0.21, r = 7$ ■

Binomial Distribution Formula

As mentioned above, the central problem of a binomial experiment is to find the probability of r successes out of n independent trials. In this section we see how to find these probabilities.

Recall from Section 5 the formula for finding the number of combinations of n distinct objects taken r at a time

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

Now, the probability of r successes in a sequence out of n independent trials is given by $p^r q^{n-r}$. Since the binomial coefficient $C(n, r)$ 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**

$$p(r) = P(X = r) = C(n, r)p^r q^{n-r}$$

where p denotes the probability of a success and $q = 1 - p$ is the probability of a failure.

Note that by letting $a = p$ and $b = 1 - p$ in the binomial formula we find

$$\sum_{k=0}^n p(k) = \sum_{k=0}^n C(n, k)p^k(1-p)^{n-k} = (p + 1 - p)^n = 1.$$

Example 18.3

Suppose that in a particular sheet of 100 postage stamps, 3 are defective. The inspection policy is to look at 5 randomly chosen stamps on a sheet and to release the sheet into circulation if none of those five is defective. Write down the random variable, the corresponding probability distribution and then determine the probability that the sheet described here will be allowed to go into circulation.

Solution.

Let X be the number of defective stamps in the sheet. Then X is a binomial random variable with probability distribution

$$P(X = x) = C(5, x)(0.03)^x(0.97)^{5-x}, \quad x = 0, 1, 2, 3, 4, 5.$$

Now,

$$P(\text{sheet goes into circulation}) = P(X = 0) = (0.97)^5 = 0.859 \blacksquare$$

Example 18.4

Suppose 40% of the student body at a large university are in favor of a ban on drinking in dormitories. Suppose 5 students 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) $P(X = 2) = C(5, 2)(.4)^2(0.6)^3 \approx 0.3456$.
- (b) $P(X < 4) = P(0) + P(1) + P(2) + P(3) = C(5, 0)(0.4)^0(0.6)^5 + C(5, 1)(0.4)^1(0.6)^4 + C(5, 2)(0.4)^2(0.6)^3 + C(5, 3)(0.4)^3(0.6)^2 \approx 0.913$.
- (c) $P(X \geq 1) = 1 - P(X < 1) = 1 - C(5, 0)(0.4)^0(0.6)^5 \approx 0.922 \blacksquare$

Example 18.5

A student has no knowledge of the material to be tested on a true-false exam with 10 questions. So, the student flips a fair coin in order to determine the response to each question.

- (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

$$\begin{aligned} P(X \geq 6) &= P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10) \\ &= \sum_{x=6}^{10} C(10, x)(0.5)^x(0.5)^{10-x} \approx 0.3769. \end{aligned}$$

(b) We have

$$P(X \leq 2) = \sum_{x=0}^2 C(10, x)(0.5)^x(0.5)^{10-x} \approx 0.0547 \blacksquare$$

Example 18.6

A pediatrician reported that 30 percent of the children in the U.S. have above normal cholesterol levels. If this is true, find the probability that in a sample of fourteen children tested for cholesterol, more than six will have above normal cholesterol levels.

Solution.

Let X be the number of children in the US with cholesterol level above the normal. Then X is a binomial random variable with $n = 14$ and $p = 0.3$. Thus, $P(X > 6) = 1 - P(X \leq 6) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) - P(X = 4) - P(X = 5) - P(X = 6) \approx 0.0933 \blacksquare$

Next, we find the expected value and variation of a binomial random variable X . 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 variation 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 18.7

The probability of hitting a target per shot is 0.2 and 10 shots are fired independently.

- (a) What is the probability that the target is hit at least twice?
- (b) What is the expected number of successful shots?
- (c) How many shots must be fired to make the probability at least 0.99 that the target will be hit?

Solution.

Let X be the number of shots which hit the target. Then, X has a binomial distribution with $n = 10$ and $p = 0.2$.

- (a) The event that the target is hit at least twice is $\{X \geq 2\}$. So,

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - P(0) - P(1) \\ &= 1 - C(10, 0)(0.2)^0(0.8)^{10} - C(10, 1)(0.2)^1(0.8)^9 \\ &\approx 0.6242 \end{aligned}$$

- (b) $E(X) = np = 10 \cdot (0.2) = 2$.

- (c) Suppose that n shots are needed to make the probability at least 0.99 that the target will be hit. Let A denote that the target is hit. Then, A^c means that the target is not hit. We have,

$$P(A) = 1 - P(A^c) = 1 - (0.8)^n \geq 0.99$$

Solving the inequality, we find that $n \geq \frac{\ln(0.01)}{\ln(0.8)} \approx 20.6$. So, the required number of shots is 21 ■

Example 18.8

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 $SD(X) = \sqrt{3}$ ■

Example 18.9

An exam consists of 25 multiple choice questions in which there are five choices for each question. Suppose that you randomly pick an answer for each question. Let X denote the total number of correctly answered questions. Write an expression that represents each of the following probabilities.

(a) The probability that you get exactly 16, or 17, or 18 of the questions correct.

(b) The probability that you get at least one of the questions correct.

Solution.

(a) $P(X = 16 \text{ or } X = 17 \text{ or } X = 18) = C(25, 16)(0.2)^{16}(0.8)^9 + C(25, 17)(0.2)^{17}(0.8)^8 + C(25, 18)(0.2)^{18}(0.8)^7$

(b) $P(X \geq 1) = 1 - P(X = 0) = 1 - C(25, 0)(0.8)^{25}$ ■

A useful fact about the binomial distribution is a recursion for calculating the probability mass function.

Theorem 18.1

Let X be a binomial random variable with parameters (n, p) . Then for $k = 1, 2, 3, \dots, n$

$$p(k) = \frac{p}{1-p} \frac{n-k+1}{k} p(k-1)$$

Proof.

We have

$$\begin{aligned} \frac{p(k)}{p(k-1)} &= \frac{C(n, k)p^k(1-p)^{n-k}}{C(n, k-1)p^{k-1}(1-p)^{n-k+1}} \\ &= \frac{\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}}{\frac{n!}{(k-1)!(n-k+1)!}p^{k-1}(1-p)^{n-k+1}} \\ &= \frac{(n-k+1)p}{k(1-p)} = \frac{p}{1-p} \frac{n-k+1}{k} \quad \blacksquare \end{aligned}$$

The following theorem details how the binomial pmf first increases and then decreases.

Theorem 18.2

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 18.10

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 18.1 ■

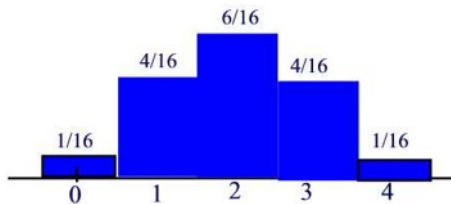


Figure 18.1

Multinomial Distribution

A **multinomial experiment** is a statistical experiment that has the following properties:

- The experiment consists of n repeated trials.
- Each trial has two or more outcomes.
- On any given trial, the probability that a particular outcome will occur is constant.
- The trials are independent; that is, the outcome on one trial does not affect the outcome on other trials.

Example 18.11

Show that the experiment of rolling two dice three times is a multinomial experiment.

Solution.

- The experiment consists of repeated trials. We toss the two dice three times.
- Each trial can result in a discrete number of outcomes - 2 through 12.
- The probability of any outcome is constant; it does not change from one toss to the next.
- The trials are independent; that is, getting a particular outcome on one trial does not affect the outcome on other trials ■

A **multinomial distribution** is the probability distribution of the outcomes from a multinomial experiment. Its probability mass function is given by the following multinomial formula:

Suppose a multinomial experiment consists of n trials, and each trial can result in any of k possible outcomes: E_1, E_2, \dots, E_k . Suppose, further, that each possible outcome can occur with probabilities p_1, p_2, \dots, p_k . Then, the probability P that E_1 occurs n_1 times, E_2 occurs n_2 times, \dots , and E_k occurs n_k times is

$$p(n_1, n_2, \dots, n_k, p_1, p_2, \dots, p_k) = \binom{n}{n_1! n_2! \dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$$

where $n = n_1 + n_2 + \dots + n_k$.

Note that

$$\sum_{\substack{(n_1, n_2, \dots, n_k) \\ n_1 + n_2 + \dots + n_k = n}} \frac{n!}{n_1! n_2! \dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k} = (p_1 + p_2 + \dots + p_k)^n = 1$$

The examples below illustrate how to use the multinomial formula to compute the probability of an outcome from a multinomial experiment.

Example 18.12

A factory makes three different kinds of bolts: Bolt A, Bolt B and Bolt C. The factory produces millions of each bolt every year, but makes twice as many of Bolt B as it does Bolt A. The number of Bolt C made is twice the total of Bolts A and B combined. Four bolts made by the factory are randomly chosen from all the bolts produced by the factory in a given year. What is the probability that the sample will contain two of Bolt B and two of Bolt C?

Solution.

Because of the proportions in which the bolts are produced, a randomly selected bolt will have a $\frac{1}{9}$ chance of being of type A, a $\frac{2}{9}$ chance of being of type B, and a $\frac{2}{3}$ chance of being of type C. A random selection of size n from the production of bolts will have a multinomial distribution with parameters n , $p_A = \frac{1}{9}$, $p_B = \frac{2}{9}$, and $p_C = \frac{2}{3}$, with probability mass function

$$P(N_A = n_A, N_B = n_B, N_C = n_C) = \frac{n!}{n_A!n_B!n_C!} \left(\frac{1}{9}\right)^{n_A} \left(\frac{2}{9}\right)^{n_B} \left(\frac{2}{3}\right)^{n_C}$$

Letting $n = 4$, $n_A = 0$, $n_B = 2$, $n_C = 2$ we find

$$P(N_A = 0, N_B = 2, N_C = 2) = \frac{4!}{0!2!2!} \left(\frac{1}{9}\right)^0 \left(\frac{2}{9}\right)^2 \left(\frac{2}{3}\right)^2 = \frac{32}{243} \blacksquare$$

Example 18.13

Suppose a card is drawn randomly from an ordinary deck of playing cards, and then put back in the deck. This exercise is repeated five times. What is the probability of drawing 1 spade, 1 heart, 1 diamond, and 2 clubs?

Solution.

To solve this problem, we apply the multinomial formula. We know the following:

- The experiment consists of 5 trials, so $n = 5$.
- The 5 trials produce 1 spade, 1 heart, 1 diamond, and 2 clubs; so $n_1 = 1, n_2 = 1, n_3 = 1$, and $n_4 = 2$.
- On any particular trial, the probability of drawing a spade, heart, diamond,

or club is 0.25, 0.25, 0.25, and 0.25, respectively. Thus, $p_1 = 0.25$, $p_2 = 0.25$, $p_3 = 0.25$, and $p_4 = 0.25$.

We plug these inputs into the multinomial formula, as shown below:

$$p(1, 1, 1, 2, 5, 0.25, 0.25, 0.25, 0.25) = \frac{5!}{1!1!1!2!} (0.25)^1 (0.25)^1 (0.25)^1 (0.25)^2 = 0.05859 \blacksquare$$

Problems

Problem 18.1

You are a telemarketer with a 10% chance of persuading a randomly selected person to switch to your long-distance company. You make 8 calls. What is the probability that exactly one is successful?

Problem 18.2

Tay-Sachs disease is a rare but fatal disease of genetic origin. If a couple are both carriers of Tay-Sachs disease, a child of theirs has probability 0.25 of being born with the disease. If such a couple has four children, what is the probability that 2 of the children have the disease?

Problem 18.3

An internet service provider (IAP) owns three servers. Each server has a 50% chance of being down, independently of the others. Let X be the number of servers 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 X 's 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 ‡ (*Trinomial Distribution*)

A large pool of adults earning their first drivers license includes 50% low-risk drivers, 30% moderate-risk drivers, and 20% high-risk drivers. Because these

drivers have no prior driving record, an insurance company considers each driver to be randomly selected from the pool.

This month, the insurance company writes 4 new policies for adults earning their first drivers license.

What is the probability that these 4 will contain at least two more high-risk drivers than low-risk drivers? Hint: Use the multinomial theorem.

Problem 18.7 ‡

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 18.8

The probability of winning a lottery is $\frac{1}{300}$. If you play the lottery 200 times, what is the probability that you win at least twice?

Problem 18.9

Suppose an airline accepted 12 reservations for a commuter plane with 10 seats. They know that 7 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 flight will be overbooked.
- (b) Find the probability that there will be empty seats.

Problem 18.10

Suppose that 3% of computer chips produced by a certain machine are defective. The chips are put into packages of 20 chips for distribution to retailers. What is the probability that a randomly selected package of chips will contain at least 2 defective chips?

Problem 18.11

The probability that a particular machine breaks down in any day is 0.20 and is independent of the breakdowns on any other day. The machine can break down only once per day. Calculate the probability that the machine breaks down two or more times in ten days.

Problem 18.12

Coins K and L are weighted so the probabilities of heads are 0.3 and 0.1, respectively. Coin K is tossed 5 times and coin L is tossed 10 times. If all the tosses are independent, what is the probability that coin K will result in heads 3 times and coin L will result in heads 6 times?

Problem 18.13

Customers at Fred's Cafe win a 100 dollar prize if their cash register receipts show a star on each of the five consecutive days Monday, \dots , Friday in any one week. The cash register is programmed to print stars on a randomly selected 10% of the receipts. If Mark eats at Fred's once each day for four consecutive weeks and the appearance of stars is an independent process, what is the standard deviation of X , where X is the number of dollars won by Mark in the four-week period?

Problem 18.14

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 18.15

Suppose we have a bowl with 10 marbles - 2 red marbles, 3 green marbles, and 5 blue marbles. We randomly select 4 marbles from the bowl, with replacement. What is the probability of selecting 2 green marbles and 2 blue marbles?

Problem 18.16

A tennis player finds that she wins against her best friend 70% of the time. They play 3 times in a particular month. Assuming independence of outcomes, what is the probability she wins at least 2 of the 3 matches?

Problem 18.17

A package of 6 fuses are tested where the probability an individual fuse is defective is 0.05. (That is, 5% of all fuses manufactured are defective).

- (a) What is the probability one fuse will be defective?
- (b) What is the probability at least one fuse will be defective?
- (c) What is the probability that more than one fuse will be defective, given at least one is defective?

Problem 18.18

In a promotion, a potato chip company inserts a coupon for a free bag in 10% of bags produced. Suppose that we buy 10 bags of chips, what is the probability that we get at least 2 coupons?

19 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

$$p(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

where λ indicates the average number of successes per unit time or space. Note that

$$\sum_{k=0}^{\infty} p(k) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

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 19.1

The number of false fire alarms in a suburb of Houston averages 2.1 per day. Assuming that a Poisson distribution is appropriate, what is the probability that 4 false alarms will occur on a given day?

Solution.

The probability that 4 false alarms will occur on a given day is given by

$$P(X = 4) = e^{-2.1} \frac{(2.1)^4}{4!} \approx 0.0992 \blacksquare$$

Example 19.2

People enter a store, on average one every two minutes.

- What is the probability that no people enter between 12:00 and 12:05?
- 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,

$$P(X = 0) = e^{-2.5} \frac{2.5^0}{0!} = e^{-2.5}.$$

(b)

$$\begin{aligned}
 P(X \geq 4) &= 1 - P(X \leq 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) \\
 &= 1 - e^{-2.5} \frac{2.5^0}{0!} - e^{-2.5} \frac{2.5^1}{1!} - e^{-2.5} \frac{2.5^2}{2!} - e^{-2.5} \frac{2.5^3}{3!} \blacksquare
 \end{aligned}$$

Example 19.3

A life insurance salesman sells on the average 3 life insurance policies per week. Use Poisson distribution to calculate the probability that in a given week he 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 he will sell one policy?

Solution.(a) Let X be the number of policies sold in a week. Then

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{e^{-3}3^0}{0!} \approx 0.95021$$

(b) We have

$$\begin{aligned}
 P(2 \leq X < 5) &= P(X = 2) + P(X = 3) + P(X = 4) \\
 &= \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!} + \frac{e^{-3}3^4}{4!} \approx 0.61611
 \end{aligned}$$

(c) Let X be the number of policies sold per day. Then $\lambda = \frac{3}{5} = 0.6$. Thus,

$$P(X = 1) = \frac{e^{-0.6}(0.6)}{1!} \approx 0.32929 \blacksquare$$

Example 19.4

A biologist suspects that the number of viruses in a small volume of blood has a Poisson distribution. She has observed the proportion of such blood samples that have no viruses is 0.01. Obtain the exact Poisson distribution she needs to use for her research. Make sure you identify the random variable first.

Solution.

Let X be the number of viruses present. Then X is a Poisson distribution with pmf

$$P(X = k) = \lambda^k \frac{e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \dots$$

Since $P(X = 0) = 0.01$ we can write $e^{-\lambda} = 0.01$. Thus, $\lambda = 4.605$ and the exact Poisson distribution is

$$P(X = k) = (4.605)^k \frac{e^{-4.605}}{k!}, \quad k = 0, 1, 2, \dots \blacksquare$$

Example 19.5

Assume that the number of defects in a certain type of magnetic tape has a Poisson distribution and that this type of tape contains, on the average, 3 defects per 1000 feet. Find the probability that a roll of tape 1200 feet long contains no defects.

Solution.

Let Y denote the number of defects on a 1200-foot long roll of tape. Then Y is a Poisson random variable with parameter $\lambda = 3 \times \frac{1200}{1000} = 3.6$. Thus,

$$P(Y = 0) = \frac{e^{-3.6}(3.6)^0}{0!} = e^{-3.6} \approx 0.0273 \blacksquare$$

The expected value of X 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 $E(X^2)$.

$$\begin{aligned}
 E(X(X-1)) &= \sum_{k=2}^{\infty} k(k-1) \frac{e^{-\lambda} \lambda^k}{k!} \\
 &= \lambda^2 \sum_{k=2}^{\infty} \frac{e^{-\lambda} \lambda^{k-2}}{(k-2)!} \\
 &= \lambda^2 \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} \\
 &= \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2
 \end{aligned}$$

Thus, $E(X^2) = E(X(X-1)) + E(X) = \lambda^2 + \lambda$ and $\text{Var}(X) = E(X^2) - (E(X))^2 = \lambda$.

Example 19.6

In the inspection of a fabric produced in continuous rolls, there is, on average, one imperfection per 10 square yards. Suppose that the number of imperfections is a random variable having Poisson distribution. Let X denote the number of imperfections in a bolt of 50 square yards. Find the mean and the variance of X .

Solution.

Since the fabric has 1 imperfection per 10 square yards, the number of imperfections in a bolt of 50 yards is 5. Thus, X is a Poisson random variable with parameter $\lambda = 5$. Hence, $E(X) = \text{Var}(X) = \lambda = 5$ ■

Poisson Approximation to the Binomial Random Variable.

Next 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 19.1

Let X be a binomial random variable with parameters n and p . If $n \rightarrow \infty$ and $p \rightarrow 0$ 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 $p \ll 1$ we can write

$$\begin{aligned} (1-p)^n &= e^{n \ln(1-p)} \\ &= e^{n(-p - \frac{p^2}{2} - \dots)} \\ &\approx e^{-np - \frac{np^2}{2}} \\ &\approx e^{-\lambda} \end{aligned}$$

We prove that

$$P(X = k) \approx e^{-\lambda} \frac{\lambda^k}{k!}.$$

This is true for $k = 0$ since $P(X = 0) = (1-p)^n \approx e^{-\lambda}$. Suppose $k > 0$ then

$$\begin{aligned} P(X = k) &= C(n, k) \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{n^k} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \\ &= \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \\ &\rightarrow 1 \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1 \end{aligned}$$

as $n \rightarrow \infty$. Note that for $0 \leq j \leq k-1$ we have $\frac{n-j}{n} = 1 - \frac{j}{n} \rightarrow 1$ as $n \rightarrow \infty$ ■

In general, Poisson distribution will provide a good approximation to binomial probabilities when $n \geq 20$ and $p \leq 0.05$. When $n \geq 100$ and $np < 10$, the approximation will generally be excellent.

Example 19.7

Let X denote the total number of people with a birthday on Christmas day in a group of 100 people. 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 Christmas day?

Solution.

We have

$$P(X \geq 1) = 1 - P(X = 0) = 1 - C(100, 0) \left(\frac{1}{365}\right)^0 \left(\frac{364}{365}\right)^{100} \approx 0.2399.$$

Using the Poisson approximation, with $\lambda = 100 \times \frac{1}{365} = \frac{100}{365} = \frac{20}{73}$ we find

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{(20/73)^0}{0!} e^{(-20/73)} \approx 0.2396 \blacksquare$$

Example 19.8

Suppose we roll two dice 12 times and we let X be the number of times a double 6 appears. Here $n = 12$ and $p = 1/36$, so $\lambda = np = 1/3$. For $k = 0, 1, 2$ compare $P(X = k)$ found using the binomial distribution with the one found using Poisson approximation.

Solution.

Let $P_B(X = k)$ be the probability using binomial distribution and $P(X = k)$ be the probability using Poisson distribution. Then

$$\begin{aligned} P_B(X = 0) &= \left(1 - \frac{1}{36}\right)^{12} \approx 0.7132 \\ P(X = 0) &= e^{-\frac{1}{3}} \approx 0.7165 \\ P_B(X = 1) &= C(12, 1) \frac{1}{36} \left(1 - \frac{1}{36}\right)^{11} \approx 0.2445 \\ P(X = 1) &= e^{-\frac{1}{3}} \frac{1}{3} \approx 0.2388 \\ P_B(X = 2) &= C(12, 2) \left(\frac{1}{36}\right)^2 \left(1 - \frac{1}{36}\right)^{10} \approx 0.0384 \\ P(X = 2) &= e^{-\frac{1}{3}} \left(\frac{1}{3}\right)^2 \frac{1}{2!} \approx 0.0398 \blacksquare \end{aligned}$$

Example 19.9

An archer shoots arrows at a circular target where the central portion of the target inside is called the bull. The archer hits the bull with probability $1/32$. Assume that the archer shoots 96 arrows at the target, and that all shoots are independent.

- Find the probability mass function of the number of bulls that the archer hits.
- Give an approximation for the probability of the archer hitting no more than one bull.

Solution.

(a) Let X denote the number of shoots that hit the bull. Then X is binomially distributed:

$$P(X = k) = C(n, k)p^k(1 - p)^{n-k}, \quad n = 96, \quad p = \frac{1}{32}.$$

(b) Since n is large, and p small, we can use the Poisson approximation, with parameter $\lambda = np = 3$. Thus,

$$P(X \leq 1) = P(X = 0) + P(X = 1) \approx e^{-\lambda} + \lambda e^{-\lambda} = 4e^{-3} \approx 0.199 \blacksquare$$

We conclude this section by establishing a recursion formula for computing $p(k)$.

Theorem 19.2

If X is a Poisson random variable with parameter λ , then

$$p(k + 1) = \frac{\lambda}{k + 1}p(k).$$

Proof.

We have

$$\begin{aligned} \frac{p(k + 1)}{p(k)} &= \frac{e^{-\lambda} \frac{\lambda^{k+1}}{(k+1)!}}{e^{-\lambda} \frac{\lambda^k}{k!}} \\ &= \frac{\lambda}{k + 1} \blacksquare \end{aligned}$$

Problems

Problem 19.1

Suppose the average number of car accidents on the highway in one day is 4. What is the probability of no car accident in one day? What is the probability of 1 car accident in two days?

Problem 19.2

Suppose the average number of calls received by an operator in one minute is 2. What is the probability of receiving 10 calls in 5 minutes?

Problem 19.3

In one particular autobiography of a professional athlete, there are an average of 15 spelling errors per page. If the Poisson distribution is used to model the probability distribution of the number of errors per page, then the random variable X , the number of errors per page, has a Poisson distribution with $\lambda = 15$. What is the probability of having no errors on a page?

Problem 19.4

Suppose that the number of accidents occurring on a highway each day is a Poisson random variable with parameter $\lambda = 3$.

- (a) Find the probability that 3 or more accidents occur today.
- (b) Find the probability that no accidents occur today.

Problem 19.5

At a checkout counter customers 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 19.6

Suppose the number of hurricanes in a given year 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 hurricanes?

Problem 19.7

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 19.8

A typesetter, on the average, makes one error in every 500 words typeset. A typical page contains 300 words. What is the probability that there will be no more than two errors in five pages?

Problem 19.9 ‡

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 19.10 ‡

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 19.11 ‡

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 19.12

The average number of trucks arriving on any one day at a truck depot in a certain city is known to be 12. What is the probability that on a given day fewer than nine trucks will arrive at this depot?

Problem 19.13

A certain kind of sheet metal has, on the average, five defects per 10 square feet. 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 19.14

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 19.15

The number of deliveries arriving at an industrial warehouse (during work hours) has a Poisson distribution with a mean of 2.5 per hour.

- (a) What is the probability an hour goes by with no more than one delivery?
- (b) Give the mean and standard deviation of the number of deliveries during 8-hour work days.

Problem 19.16

Imperfections in rolls of fabric follow a Poisson process with a mean of 64 per 10000 feet. Give the approximate probability a roll will have at least 75 imperfections.

20 Other Discrete Random Variables

20.1 Geometric Random Variable

A **geometric** random variable with parameter $p, 0 < p < 1$ has a probability mass function

$$p(n) = p(1 - p)^{n-1}, \quad n = 1, 2, \dots$$

A geometric random variable models the number of successive 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 20.1

If you roll a pair of fair dice, the probability of getting an 11 is $\frac{1}{18}$. If you roll the dice repeatedly, what is the probability that the first 11 occurs on the 8th roll?

Solution.

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,

$$P(X = 8) = \left(\frac{1}{18}\right) \left(1 - \frac{1}{18}\right)^7 = 0.0372 \blacksquare$$

To find the expected value and variance of a geometric random variable we proceed as follows. First we recall that for $0 < x < 1$ we have $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$. Differentiating this twice we find $\sum_{n=1}^{\infty} nx^{n-1} = (1-x)^{-2}$ and $\sum_{n=1}^{\infty} n(n-1)x^{n-2} = 2(1-x)^{-3}$. Letting $x = 1-p$ in the last two equations we find

$$\sum_{n=1}^{\infty} n(1-p)^{n-1} = p^{-2} \quad \text{and} \quad \sum_{n=1}^{\infty} n(n-1)(1-p)^{n-2} = 2p^{-3}.$$

We next apply these equalities in finding $E(X)$ and $E(X^2)$. Indeed, we have

$$\begin{aligned} E(X) &= \sum_{n=1}^{\infty} n(1-p)^{n-1}p \\ &= p \sum_{n=1}^{\infty} n(1-p)^{n-1} \\ &= p \cdot p^{-2} = p^{-1} \end{aligned}$$

and

$$\begin{aligned} E(X(X-1)) &= \sum_{n=1}^{\infty} n(n-1)(1-p)^{n-1}p \\ &= p(1-p) \sum_{n=1}^{\infty} n(n-1)(1-p)^{n-2} \\ &= p(1-p) \cdot (2p^{-3}) = 2p^{-2}(1-p) \end{aligned}$$

so that $E(X^2) = (2-p)p^{-2}$. The variance is then given by

$$\text{Var}(X) = E(X^2) - (E(X))^2 = (2-p)p^{-2} - p^{-2} = \frac{1-p}{p^2}.$$

Note that

$$\sum_{n=1}^{\infty} p(1-p)^{n-1} = \frac{p}{1-(1-p)} = 1.$$

Also, observe that for $k = 1, 2, \dots$ we have

$$P(X \geq k) = \sum_{n=k}^{\infty} p(1-p)^{n-1} = p(1-p)^{k-1} \sum_{n=0}^{\infty} (1-p)^n = \frac{p(1-p)^{k-1}}{1-(1-p)} = (1-p)^{k-1}$$

and

$$P(X \leq k) = 1 - P(X \geq k+1) = 1 - (1-p)^k.$$

From this, one can find the cdf given by

$$F(X) = P(X \leq x) = \begin{cases} 0 & x < 1 \\ 1 - (1-p)^k & k \leq x < k+1, \quad k = 1, 2, \dots \end{cases}$$

We can use $F(x)$ for computing probabilities such as $P(j \leq X \leq k)$, which is the probability that it will take from j attempts to k attempts to succeed.

This value is given by

$$\begin{aligned} P(j \leq X \leq k) &= P(X \leq k) - P(X \leq j-1) = (1 - (1-p)^k) - (1 - (1-p)^{j-1}) \\ &= (1-p)^{j-1} - (1-p)^k \end{aligned}$$

Remark 20.1

The fact that $P(j \leq X \leq k) = P(X \leq k) - P(X \leq j-1)$ follows from Proposition 21.9 of Section 21.

Example 20.2

Computer chips are tested one at a time until a good chip is found. Let X denote the number of chips that need to be tested in order to find a good one. Given that $P(X > 3) = 1/2$, what is $E(X)$?

Solution.

X has geometric distribution, so $P(X > 3) = P(X \geq 4) = (1 - p)^3$. Setting this equal to $1/2$ and solving for p gives $p = 1 - 2^{-\frac{1}{3}}$. Therefore,

$$E(X) = \frac{1}{p} = \frac{1}{1 - 2^{-\frac{1}{3}}} \blacksquare$$

Example 20.3

From past experience it is known that 3% of accounts in a large accounting population are in error.

- (a) What is the probability that 5 accounts are audited before an account in error is found?
 (b) What is the probability that the first account in error occurs in the first five accounts audited?

Solution.

(a) Let X be the number of inspections to obtain first account in error. Then X is a geometric random variable with $p = 0.03$. Thus

$$P(X = 5) = (0.03)(0.97)^4.$$

(b)

$$P(X \leq 5) = 1 - P(X \geq 6) = 1 - (0.97)^5 \approx 0.141 \blacksquare$$

Example 20.4

Suppose you keep rolling four dice simultaneously until at least one of them shows a 6. What is the expected number of “rounds” (each round consisting of a simultaneous roll of four dice) you have to play?

Solution.

If “success” of a round is defined as “at least one 6” then X is simply the number of trials needed to get the first success, with p given by $p = 1 - \left(\frac{5}{6}\right)^4 \approx 0.51775$. Thus, X has geometric distribution, and so $E(X) = \frac{1}{p} \approx 1.9314 \blacksquare$

Example 20.5

Assume that every time you attend your 2027 lecture there is a probability of 0.1 that your Professor will not show up. Assume her arrival to any given lecture is independent of her arrival (or non-arrival) to any other lecture. What is the expected number of classes you must attend until you arrive to find your Professor absent?

Solution.

Let X be the number of classes you must attend until you arrive to find your Professor absent, then X has Geometric distribution with parameter $p = 0.1$. Thus

$$P(X = n) = 0.1(1 - p)^{n-1}, \quad n = 1, 2, \dots$$

and

$$E(X) = \frac{1}{p} = 10 \blacksquare$$

Problems

Problem 20.1

An urn contains 5 white, 4 black, and 1 red marble. Marbles are drawn, with replacement, until a red one is found. If X is the random variable counting the number of trials until a red marble appears, then

- (a) What is the probability that the red marble appears on the first trial?
- (b) What is the probability that the red marble appears on the second trial?
- (c) What is the probability that the marble appears on the k^{th} trial.

Problem 20.2

The probability that a machine produces a defective item is 0.10. Each item is checked as it is produced. Assume that these are independent trials, find the probability that at least 10 items must be checked to find one that is defective.

Problem 20.3

Suppose parts are of two varieties: good (with probability $90/92$) and slightly defective (with probability $2/92$). Parts are produced one after the other. What is the probability that at least 5 parts must be produced until there is a slightly defective part produced?

Problem 20.4

Assume that every time you drive your car, there is a 0.001 probability that you will receive a speeding ticket, independent of all other times you have driven.

- (a) What is the probability you will drive your car two or less times before receiving your first ticket?
- (b) What is the expected number of times you will drive your car until you receive your first ticket?

Problem 20.5

Three people study together. They don't like to ask questions in class, so they agree that they will flip a fair coin and if one of them has a different outcome than the other two, that person will ask the group's question in class. If all three match, they flip again until someone is the questioner. What is the probability that

- (a) exactly three rounds of flips are needed?
- (b) more than four rounds are needed?

Problem 20.6

Suppose one die is rolled over and over until a Two is rolled. What is the probability that it takes from 3 to 6 rolls?

Problem 20.7

You have a fair die that you roll over and over until you get a 5 or a 6. (You stop rolling it once it shows a 5 or a 6.) Let X be the number of times you roll the die.

- (a) What is $P(X = 3)$? What is $P(X = 50)$?
- (b) What is $E(X)$?

Problem 20.8

Fifteen percent of houses in your area have finished basements. Your real estate agent starts showing you homes at random, one after the other. Let X be the number of homes with finished basements that you see, before the first house that has no finished basement.

- (a) What is the probability distribution of X ?
- (b) What is the probability distribution of $Y = X + 1$?

Problem 20.9

Suppose that 3% of computer chips produced by a certain machine are defective. The chips are put into packages of 20 chips for distribution to retailers.

- (a) What is the probability that a randomly selected package of chips will contain at least 2 defective chips?
- (b) Suppose we continue to select packs of bulbs randomly from the production. What is the probability that it will take fewer than five packs to find a pack with at least 2 defective chips?

Problem 20.10

Show that the Geometric distribution with parameter p satisfies the equation

$$P(X > i + j | X > i) = P(X > j).$$

This says that the Geometric distribution satisfies the **memoryless property**

Problem 20.11 ‡

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 20.12

An actress has a probability of getting offered a job after a try-out of $p = 0.10$. She plans to keep trying out for new jobs until she gets offered. Assume outcomes of try-outs are independent.

- (a) How many try-outs does she expect to have to take?
- (b) What is the probability she will need to attend more than 2 try-outs?

Problem 20.13

Which of the following sampling plans is binomial and which is geometric? Give $P(X = x) = p(x)$ for each case. Assume outcomes of individual trials are independent with constant probability of success (e.g. Bernoulli trials).

- (a) A missile designer will launch missiles until one successfully reaches its target. The underlying probability of success is 0.40.
- (b) A clinical trial enrolls 20 children with a rare disease. Each child is given an experimental therapy, and the number of children showing marked improvement is observed. The true underlying probability of success is 0.60.

20.2 Negative Binomial Random Variable

Consider a statistical experiment where a success occurs with probability p and a failure occurs with probability $q = 1 - p$. If the experiment is repeated indefinitely and the trials are independent of each other, then the random variable X , the number of trials at which the r^{th} success occurs, has a **negative binomial** distribution with parameters r and p . The probability mass function of X is

$$p(n) = P(X = n) = C(n - 1, r - 1)p^r(1 - p)^{n-r},$$

where $n = r, r + 1, \dots$ (In order to have r successes there must be at least r trials.)

For the r^{th} success to occur on the n^{th} trial, there must have been $r - 1$ successes and $n - r$ failures among the first $n - 1$ trials. The number of ways of distributing $r - 1$ successes among $n - 1$ trials is $C(n - 1, r - 1)$. 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 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. 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

$$P(Y = y) = C(r + y - 1, y)p^r(1 - p)^y, \quad y = 0, 1, 2, \dots$$

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 $C(r + y - 1, y) = C(r + y - 1, r - 1)$.

The negative binomial distribution gets its name from the relationship

$$C(r + y - 1, y) = (-1)^y \frac{(-r)(-r - 1) \cdots (-r - y + 1)}{y!} = (-1)^y C(-r, y)$$

which is the defining equation for binomial coefficient with negative integers. Now, recalling the binomial series expansion

$$\begin{aligned}(1-t)^{-r} &= \sum_{k=0}^{\infty} (-1)^k C(-r, k) t^k \\ &= \sum_{k=0}^{\infty} C(r+k-1, k) t^k, \quad -1 < t < 1\end{aligned}$$

Thus,

$$\begin{aligned}\sum_{y=0}^{\infty} P(Y=y) &= \sum_{y=0}^{\infty} C(r+y-1, y) p^r (1-p)^y \\ &= p^r \sum_{y=0}^{\infty} C(r+y-1, y) (1-p)^y \\ &= p^r \cdot p^{-r} = 1\end{aligned}$$

This shows that $p(n)$ is indeed a probability mass function.

Example 20.6

A research scientist is inoculating rabbits, one at a time, with a disease until he finds two rabbits which develop the disease. If the probability of contracting the disease $\frac{1}{6}$, what is the probability that eight rabbits are needed?

Solution.

Let X be the number of rabbits needed until the first rabbit to contract the disease. Then X follows a negative binomial distribution with $r = 2$, $n = 6$, and $p = \frac{1}{6}$. Thus,

$$P(8 \text{ rabbits are needed}) = C(2-1+6, 6) \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^6 \approx 0.0651 \blacksquare$$

Example 20.7

Suppose we are at a rifle range with an old gun that misfires 5 out of 6 times. Define “success” as the event the gun fires 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?

Solution.

The probability that there are 10 failures before the third success is given by

$$P(Y = 10) = C(12, 10) \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^{10} \approx 0.0493 \blacksquare$$

Example 20.8

A comedian tells jokes that are only funny 25% of the time. What is the probability that he tells his tenth funny joke in his fortieth attempt?

Solution.

A success is when a joke is funny. Let X number of attempts at which the tenth joke occurs. Then X is a negative binomial random variable with parameters $r = 10$ and $p = 0.25$. Thus,

$$P(X = 40) = C(40 - 1, 10 - 1)(0.25)^{10}(0.75)^{30} \approx 0.0360911 \blacksquare$$

The expected value of Y is

$$\begin{aligned} E(Y) &= \sum_{y=0}^{\infty} yC(r + y - 1, y)p^r(1 - p)^y \\ &= \sum_{y=1}^{\infty} \frac{(r + y - 1)!}{(y - 1)!(r - 1)!} p^r(1 - p)^y \\ &= \sum_{y=1}^{\infty} \frac{r(1 - p)}{p} C(r + y - 1, y - 1)p^{r+1}(1 - p)^{y-1} \\ &= \frac{r(1 - p)}{p} \sum_{z=0}^{\infty} C((r + 1) + z - 1, z)p^{r+1}(1 - p)^z \\ &= \frac{r(1 - p)}{p} \end{aligned}$$

It follows that

$$E(X) = E(Y + r) = E(Y) + r = \frac{r}{p}.$$

Similarly,

$$\begin{aligned}
 E(Y^2) &= \sum_{y=0}^{\infty} y^2 C(r+y-1, y) p^r (1-p)^y \\
 &= \frac{r(1-p)}{p} \sum_{y=1}^{\infty} y \frac{(r+y-1)!}{(y-1)! r!} p^{r+1} (1-p)^{y-1} \\
 &= \frac{r(1-p)}{p} \sum_{z=0}^{\infty} (z+1) C((r+1)+z-1, z) p^{r+1} (1-p)^z \\
 &= \frac{r(1-p)}{p} (E(Z) + 1)
 \end{aligned}$$

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 $X = Y + r$ then

$$\text{Var}(X) = \text{Var}(Y) = \frac{r(1-p)}{p^2}.$$

Example 20.9

Suppose we are at a rifle range with an old gun that misfires 5 out of 6 times. Define "success" as the event the gun fires and let X be the number of failures before the third success. Then X is a negative binomial random variable with parameters $(3, \frac{1}{6})$. Find $E(X)$ and $\text{Var}(X)$.

Solution.

The expected value of X is

$$E(X) = \frac{r(1-p)}{p} = 15$$

and the variance is

$$\text{Var}(X) = \frac{r(1-p)}{p^2} = 90 \blacksquare$$

Problems

Problem 20.14

A phenomenal major-league baseball player has a batting average of 0.400. Beginning with his next at-bat, the random variable X , whose value refers to the number of the at-bat (walks, sacrifice flies and certain types of outs are not considered at-bats) when his r^{th} hit occurs, has a negative binomial distribution with parameters r and $p = 0.400$. What is the probability that this hitter's second hit comes on the fourth at-bat?

Problem 20.15

A geological study indicates that an exploratory oil well drilled in a particular region should strike oil with $p = 0.2$. Find $P(\text{the } 3^{\text{rd}} \text{ oil strike comes on the } 5^{\text{th}} \text{ well drilled})$.

Problem 20.16

We randomly draw a card with replacement from a 52-card deck and record its face value and then put it back. Our goal is to successfully choose three aces. Let X be the number of trials we need to perform in order to achieve our goal.

- (a) What is the distribution of X ?
- (b) What is the probability that $X = 39$?

Problem 20.17

Find the expected value and the variance of the number of times one must throw a die until the outcome 1 has occurred 4 times.

Problem 20.18

Find the probability that a man flipping a coin gets the fourth head on the ninth flip.

Problem 20.19

The probability that a driver passes the written test for a driver's license is 0.75. What is the probability that a person will fail the test on the first try and pass the test on the second try?

Problem 20.20 ‡

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 20.21

Waiters are extremely distracted today and have a 0.6 chance of making a mistake with your coffee order, giving you decaf even though you order caffeinated. Find the probability that your order of the second decaffeinated coffee occurs on the seventh order of regular coffees.

Problem 20.22

Suppose that 3% of computer chips produced by a certain machine are defective. The chips are put into packages of 20 chips for distribution to retailers.

(a) What is the probability that a randomly selected package of chips will contain at least 2 defective chips?

(b) What is the probability that the tenth pack selected is the third to contain at least two defective chips?

Problem 20.23

Assume that every time you attend your 2027 lecture there is a probability of 0.1 that your Professor will not show up. Assume her arrival to any given lecture is independent of her arrival (or non-arrival) to any other lecture. What is the expected number of classes you must attend until the second time you arrive to find your Professor absent?

Problem 20.24

If the probability is 0.40 that a child exposed to a certain contagious disease will catch it, what is the probability that the tenth child exposed to the disease will be the third to catch it?

Problem 20.25

In rolling a fair die repeatedly (and independently on successive rolls), find the probability of getting the third "1" on the t -th roll.

20.3 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? There are $C(N, r)$ r -subsets of the group. There are $C(n, k)$ k -subsets of the men and $C(N - n, r - k)$ $(r - k)$ -subsets of the women. Thus the probability of getting exactly k men on the committee is

$$p(k) = P(X = k) = \frac{C(n, k)C(N - n, r - k)}{C(N, r)}, \quad k = 0, 1, \dots, r, \quad r < \min\{n, N - n\}$$

This is the probability mass function of X . Note that

$$\sum_{k=0}^r \frac{C(n, k)C(N - n, r - k)}{C(N, r)} = 1$$

The proof of this result follows from **Vandermonde's identity**

Theorem 20.1

$$C(n + m, r) = \sum_{k=0}^r C(n, k)C(m, r - k)$$

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 $C(n + m, r)$. 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 20.10

Suppose a sack contains 70 red beads and 30 green ones. If we draw out 20 without replacement, what is the probability of getting exactly 14 red ones?

Solution.

If X is the number of red beads, then X is a hypergeometric random variable with parameters $N = 100, r = 20, n = 70$. Thus,

$$P(X = 14) = \frac{C(70, 14)C(30, 6)}{C(100, 20)} \approx 0.21 \blacksquare$$

Example 20.11

13 Democrats, 12 Republicans and 8 Independents are sitting in a room. 8 of these people will be selected to serve on a special committee. What is the probability that exactly 5 of the committee members will be Democrats?

Solution.

Let X be the number of democrats in the committee. Then X is hypergeometric random variable with parameters $N = 33, r = 8, n = 13$. Thus,

$$P(X = 5) = \frac{C(13, 5)C(20, 3)}{C(33, 8)} \approx 0.00556 \blacksquare$$

Next, we find the expected value of a hypergeometric random variable with parameters N, n, r . Let k be a positive integer. Then

$$\begin{aligned} E(X^k) &= \sum_{i=0}^r i^k P(X = i) \\ &= \sum_{i=0}^r i^k \frac{C(n, i)C(N - r, r - i)}{C(N, r)} \end{aligned}$$

Using the identities

$$iC(n, i) = nC(n - 1, i - 1) \text{ and } rC(N, r) = NC(N - 1, r - 1)$$

we obtain that

$$\begin{aligned} E(X^k) &= \frac{nr}{N} \sum_{i=1}^r i^{k-1} \frac{C(n-1, i-1)C(N-r, r-i)}{C(N-1, r-1)} \\ &= \frac{nr}{N} \sum_{j=0}^{r-1} (j+1)^{k-1} \frac{C(n-1, j)C((N-1)-(r-1), (r-1)-j)}{C(N-1, r-1)} \\ &= \frac{nr}{N} E[(Y+1)^{k-1}] \end{aligned}$$

where Y is a hypergeometric random variable with parameters $N-1$, $n-1$, 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,

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

Example 20.12

The United States senate has 100 members. Suppose there are 54 Republicans and 46 Democrats. A committee of 15 senators is selected at random.

- What is the probability that there will be 9 Republicans and 6 Democrats on the committee?
- What is the expected number of Republicans on this committee?
- What is the variance of the number of Republicans on this committee?

Solution.

Let X be the number of Republicans of the committee of 15 selected at random. Then X is a hypergeometric random variable with $N=100$, $r=54$, and $n=15$.

- $P(X=9) = \frac{C(54,9)C(46,6)}{C(100,15)}$.
- $E(X) = \frac{nr}{N} = 15 \times \frac{54}{100} = 8.1$
- $\text{Var}(X) = n \cdot \frac{r}{N} \cdot \frac{N-r}{N} \cdot \frac{N-n}{N-1} = 3.199$ ■

Example 20.13

A lot of 15 semiconductor chips contains 6 defective chips and 9 good chips. Five chips are randomly selected without replacement.

- (a) What is the probability that there are 2 defective and 3 good chips in the sample?
(b) What is the probability that there are at least 3 good 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 $r = 6$, $N = 15$, $n = 5$. The desired probability is

$$P(X = 2) = \frac{C(6, 2)C(9, 3)}{C(15, 5)} = \frac{420}{1001}$$

(b) Note that the event that there are at least 3 good 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

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{C(6, 0)C(9, 5)}{C(15, 5)} + \frac{C(6, 1)C(9, 4)}{C(15, 5)} + \frac{C(6, 2)C(9, 3)}{C(15, 5)} \\ &= \frac{714}{1001} \end{aligned}$$

(c) $E(X) = \frac{rN}{n} = 5 \cdot \frac{6}{15} = 2$ ■

Problems

Problem 20.26

Suppose we randomly select 5 cards without replacement from an ordinary deck of playing cards. What is the probability of getting exactly 2 red cards (i.e., hearts or diamonds)?

Problem 20.27

There are 25 socks in a box, 15 red and 10 blue. Pick 7 without replacement. What is the probability of picking exactly 3 red socks?

Problem 20.28

A player in the California lottery chooses 6 numbers from 53 and the lottery officials later choose 6 numbers at random. Let X equal the number of matches. Find the probability distribution function.

Problem 20.29

Compute the probability of obtaining 3 defectives in a sample of size 10 without replacement from a box of 20 components containing 4 defectives.

Problem 20.30

A bag contains 10 \$50 bills and 190 \$1 bills, and 10 bills are drawn without replacement, what is the probability that 2 of the bills drawn are \$50s?

Problem 20.31

A package of 8 AA batteries contains 2 batteries that are defective. A student randomly selects four batteries and replaces the batteries in his calculator.

- (a) What is the probability that all four batteries work?
- (b) What are the mean and variance for the number of batteries that work?

Problem 20.32

Suppose a company fleet of 20 cars contains 7 cars that do not meet government exhaust emissions standards and are therefore releasing excessive pollution. Moreover, suppose that a traffic policeman randomly inspects 5 cars. What is the probability of no more than 2 polluting cars being selected?

Problem 20.33

There are 123,850 pickup trucks in Austin. Of these, 2,477 are stolen. Suppose that 100 randomly chosen pickups are checked by the police. What is the probability that exactly 3 of the 100 chosen pickups are stolen? Give the expression without finding the exact numeric value.

Problem 20.34

Consider an urn with 7 red balls and 3 blue balls. Suppose we draw 4 balls without replacement and let X be the total number of red balls we get. Compute $P(X \leq 1)$.

Problem 20.35

As part of an air-pollution survey, an inspector decided to examine the exhaust of six of a company's 24 trucks. If four of the company's trucks emit excessive amounts of pollutants, what is the probability that none of them will be included in the inspector's sample?

Problem 20.36

A fair die is tossed until a 2 is obtained. If X is the number of trials required to obtain the first 2, what is the smallest value of x for which $P(X \leq x) \geq \frac{1}{2}$?

Problem 20.37

A box contains 10 white and 15 black marbles. Let X denote the number of white marbles in a selection of 10 marbles selected at random and without replacement. Find $\frac{\text{Var}(X)}{E(X)}$.

Problem 20.38

Among the 48 applicants for a job, 30 have college degrees. 10 of the applicants are randomly chosen for interviews. Let X be the number of applicants among these ten who have college degrees. Find $P(X \leq 8)$.

21 Properties of the Cumulative Distribution Function

Random variables are classified into three types rather than two: Continuous, discrete, and mixed.

A random variable will be called a continuous type random variable if its cumulative distribution function is continuous. As a matter of fact, its cumulative distribution function is a continuous nondecreasing function. Thus, its cumulative distribution graph has no jumps. A thorough discussion about this type of random variables starts in the next section.

On the other extreme are the discrete type random variables, which are all about jumps as you will notice later on in this section. Indeed, the graph of the cumulative distribution function is a step function.

A random variable whose cumulative distribution function is partly discrete and partly continuous is called a **mixed** random variable.

In this section, we will discuss properties of the cumulative distribution function that are valid to all three types.

Recall from Section 14 that if X is a random variable (discrete or continuous) then the **cumulative distribution function** (abbreviated c.d.f) is the function

$$F(t) = P(X \leq t).$$

In this section we discuss some of the properties of c.d.f and its applications. 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

$$\lim_{n \rightarrow \infty} E_n = \bigcup_{n=1}^{\infty} E_n.$$

For a decreasing sequence we define

$$\lim_{n \rightarrow \infty} E_n = \bigcap_{n=1}^{\infty} E_n.$$

Proposition 21.1

If $\{E_n\}_{n \geq 1}$ is either an increasing or decreasing sequence of events then

(a)

$$\lim_{n \rightarrow \infty} P(E_n) = P(\lim_{n \rightarrow \infty} E_n)$$

that is

$$P(\cup_{n=1}^{\infty} E_n) = \lim_{n \rightarrow \infty} P(E_n)$$

for increasing sequence and

(b)

$$P(\cap_{n=1}^{\infty} E_n) = \lim_{n \rightarrow \infty} P(E_n)$$

for decreasing sequence.

Proof.

(a) Suppose first that $E_n \subset E_{n+1}$ for all $n \geq 1$. Define the events

$$\begin{aligned} F_1 &= E_1 \\ F_n &= E_n \cap E_{n-1}^c, \quad n > 1 \end{aligned}$$

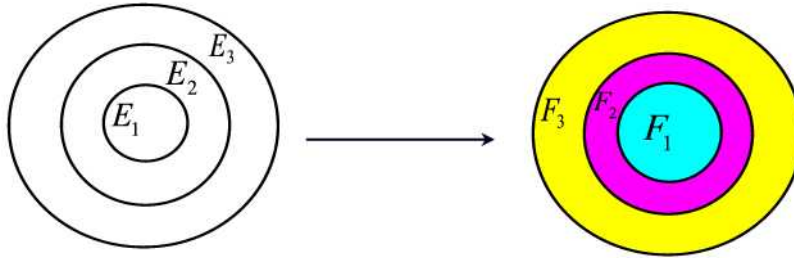


Figure 21.1

These events are shown in the Venn diagram of Figure 21.1. 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, $\cup_{n=1}^{\infty} F_n = \cup_{n=1}^{\infty} E_n$

and for $n \geq 1$ we have $\cup_{i=1}^n F_i = \cup_{i=1}^n E_i$. From these properties we have

$$\begin{aligned}
 P(\lim_{n \rightarrow \infty} E_n) &= P(\cup_{n=1}^{\infty} E_n) \\
 &= P(\cup_{n=1}^{\infty} F_n) \\
 &= \sum_{n=1}^{\infty} P(F_n) \\
 &= \lim_{n \rightarrow \infty} \sum_{i=1}^n P(F_i) \\
 &= \lim_{n \rightarrow \infty} P(\cup_{i=1}^n F_i) \\
 &= \lim_{n \rightarrow \infty} P(\cup_{i=1}^n E_i) \\
 &= \lim_{n \rightarrow \infty} P(E_n)
 \end{aligned}$$

(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

$$P(\cup_{n=1}^{\infty} E_n^c) = \lim_{n \rightarrow \infty} P(E_n^c)$$

By De Morgan's Law we have $\cup_{n=1}^{\infty} E_n^c = (\cap_{n=1}^{\infty} E_n)^c$. Thus,

$$P((\cap_{n=1}^{\infty} E_n)^c) = \lim_{n \rightarrow \infty} P(E_n^c).$$

Equivalently,

$$1 - P(\cap_{n=1}^{\infty} E_n) = \lim_{n \rightarrow \infty} [1 - P(E_n)] = 1 - \lim_{n \rightarrow \infty} P(E_n)$$

or

$$P(\cap_{n=1}^{\infty} E_n) = \lim_{n \rightarrow \infty} P(E_n) \blacksquare$$

Proposition 21.2

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 $P(X \leq a) \leq P(X \leq b)$. Hence, $F(a) \leq F(b)$ ■

Example 21.1

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.

No because $F(2) < F(1)$ ■

Proposition 21.3

F is continuous from the right. That is, $\lim_{t \rightarrow b^+} F(t) = F(b)$.

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 21.1 we have

$$\lim_{n \rightarrow \infty} F(b_n) = \lim_{n \rightarrow \infty} P(E_n) = P(\bigcap_{n=1}^{\infty} E_n) = P(X \leq b) = F(b) \blacksquare$$

Proposition 21.4

(a) $\lim_{b \rightarrow -\infty} F(b) = 0$

(b) $\lim_{b \rightarrow \infty} F(b) = 1$

Proof.

(a) Let $\{b_n\}_{n \geq 1}$ be a decreasing sequence with $\lim_{n \rightarrow \infty} b_n = -\infty$. Define $E_n = \{s : X(s) \leq b_n\}$. Then $\{E_n\}_{n \geq 1}$ is a decreasing sequence of events such that $\bigcap_{n=1}^{\infty} E_n = \{s : X(s) < -\infty\}$. By Proposition 21.1 we have

$$\lim_{n \rightarrow \infty} F(b_n) = \lim_{n \rightarrow \infty} P(E_n) = P(\bigcap_{n=1}^{\infty} E_n) = P(X < -\infty) = 0.$$

(b) Let $\{b_n\}_{n \geq 1}$ be an increasing sequence with $\lim_{n \rightarrow \infty} b_n = \infty$. Define $E_n = \{s : X(s) \leq b_n\}$. Then $\{E_n\}_{n \geq 1}$ is an increasing sequence of events such that $\bigcup_{n=1}^{\infty} E_n = \{s : X(s) < \infty\}$. By Proposition 21.1 we have

$$\lim_{n \rightarrow \infty} F(b_n) = \lim_{n \rightarrow \infty} P(E_n) = P(\bigcup_{n=1}^{\infty} E_n) = P(X < \infty) = 1 \blacksquare$$

Example 21.2

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 ■

All probability questions can be answered in terms of the c.d.f.

Proposition 21.5

$$P(X > a) = 1 - F(a).$$

Proof.

We have $P(X > a) = 1 - P(X \leq a) = 1 - F(a)$ ■

Example 21.3

Let X have probability mass function (pmf) $p(x) = \frac{1}{8}$ for $x = 1, 2, \dots, 8$.

Find

- (a) the cumulative distribution function (cdf) of X ;
- (b) $P(X > 5)$.

Solution.

(a) The cdf is given by

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{[x]}{8} & 1 \leq x \leq 8 \\ 1 & x > 8 \end{cases}$$

where $[x]$ is the integer part of x .

(b) We have $P(X > 5) = 1 - F(5) = 1 - \frac{5}{8} = \frac{3}{8}$ ■

Proposition 21.6

$$P(X < a) = \lim_{n \rightarrow \infty} F\left(a - \frac{1}{n}\right) = F(a^-).$$

Proof.

We have

$$\begin{aligned} P(X < a) &= P\left(\lim_{n \rightarrow \infty} \left\{X \leq a - \frac{1}{n}\right\}\right) \\ &= \lim_{n \rightarrow \infty} P\left(X \leq a - \frac{1}{n}\right) \\ &= \lim_{n \rightarrow \infty} F\left(a - \frac{1}{n}\right) \quad \blacksquare \end{aligned}$$

Note that $P(X < a)$ does not necessarily equal $F(a)$, since $F(a)$ also includes the probability that X equals a .

Corollary 21.1

$$P(X \geq a) = 1 - \lim_{n \rightarrow \infty} F\left(a - \frac{1}{n}\right) = 1 - F(a^-).$$

Proposition 21.7

If $a < b$ then $P(a < X \leq b) = F(b) - F(a)$.

Proof.

Let $A = \{s : X(s) > a\}$ and $B = \{s : X(s) \leq b\}$. Note that $P(A \cup B) = 1$. Then

$$\begin{aligned} P(a < X \leq b) &= P(A \cap B) \\ &= P(A) + P(B) - P(A \cup B) \\ &= (1 - F(a)) + F(b) - 1 = F(b) - F(a) \blacksquare \end{aligned}$$

Proposition 21.8

If $a < b$ then $P(a \leq X < b) = \lim_{n \rightarrow \infty} F\left(b - \frac{1}{n}\right) - \lim_{n \rightarrow \infty} F\left(a - \frac{1}{n}\right)$.

Proof.

Let $A = \{s : X(s) \geq a\}$ and $B = \{s : X(s) < b\}$. Note that $P(A \cup B) = 1$. Then using Proposition 21.6 we find

$$\begin{aligned} P(a \leq X < b) &= P(A \cap B) \\ &= P(A) + P(B) - P(A \cup B) \\ &= \left(1 - \lim_{n \rightarrow \infty} F\left(a - \frac{1}{n}\right)\right) + \lim_{n \rightarrow \infty} F\left(b - \frac{1}{n}\right) - 1 \\ &= \lim_{n \rightarrow \infty} F\left(b - \frac{1}{n}\right) - \lim_{n \rightarrow \infty} F\left(a - \frac{1}{n}\right) \blacksquare \end{aligned}$$

Proposition 21.9

If $a < b$ then $P(a \leq X \leq b) = F(b) - \lim_{n \rightarrow \infty} F\left(a - \frac{1}{n}\right) = F(b) - F(a^-)$.

Proof.

Let $A = \{s : X(s) \geq a\}$ and $B = \{s : X(s) \leq b\}$. Note that $P(A \cup B) = 1$. Then

$$\begin{aligned} P(a \leq X \leq b) &= P(A \cap B) \\ &= P(A) + P(B) - P(A \cup B) \\ &= \left(1 - \lim_{n \rightarrow \infty} F\left(a - \frac{1}{n}\right)\right) + F(b) - 1 \\ &= F(b) - \lim_{n \rightarrow \infty} F\left(a - \frac{1}{n}\right) \blacksquare \end{aligned}$$

Example 21.4

Show that $P(X = a) = F(a) - F(a^-)$.

Solution.

Applying the previous result we can write $P(a \leq x \leq a) = F(a) - F(a^-)$ ■

Proposition 21.10

If $a < b$ then $P(a < X < b) = \lim_{n \rightarrow \infty} F\left(b - \frac{1}{n}\right) - F(a)$.

Proof.

Let $A = \{s : X(s) > a\}$ and $B = \{s : X(s) < b\}$. Note that $P(A \cup B) = 1$. Then

$$\begin{aligned} P(a < X < b) &= P(A \cap B) \\ &= P(A) + P(B) - P(A \cup B) \\ &= (1 - F(a)) + \lim_{n \rightarrow \infty} F\left(b - \frac{1}{n}\right) - 1 \\ &= \lim_{n \rightarrow \infty} F\left(b - \frac{1}{n}\right) - F(a) \blacksquare \end{aligned}$$

Figure 21.2 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, \dots is equal to the probability that X assumes that particular value.

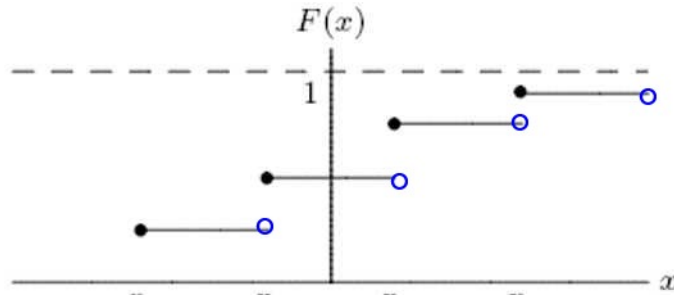


Figure 21.2

Example 21.5 (Mixed RV)

The distribution function of a random variable X , is given by

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{2}, & 0 \leq x < 1 \\ \frac{2}{3}, & 1 \leq x < 2 \\ \frac{11}{12}, & 2 \leq x < 3 \\ 1, & 3 \leq x \end{cases}$$

- Graph $F(x)$.
- Compute $P(X < 3)$.
- Compute $P(X = 1)$.
- Compute $P(X > \frac{1}{2})$.
- Compute $P(2 < X \leq 4)$.

Solution.

(a) The graph is given in Figure 21.3.

(b) $P(X < 3) = \lim_{n \rightarrow \infty} P\left(\left\{X \leq 3 - \frac{1}{n}\right\}\right) = \lim_{n \rightarrow \infty} F\left(3 - \frac{1}{n}\right) = \frac{11}{12}$.

(c) $P(X = 1) = P(X \leq 1) - P(X < 1) = F(1) - \lim_{n \rightarrow \infty} F\left(1 - \frac{1}{n}\right) = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$.

(d) $P(X > \frac{1}{2}) = 1 - P(X \leq \frac{1}{2}) = 1 - F(\frac{1}{2}) = \frac{3}{4}$.

(e) $P(2 < X \leq 4) = F(4) - F(2) = \frac{1}{12}$ ■

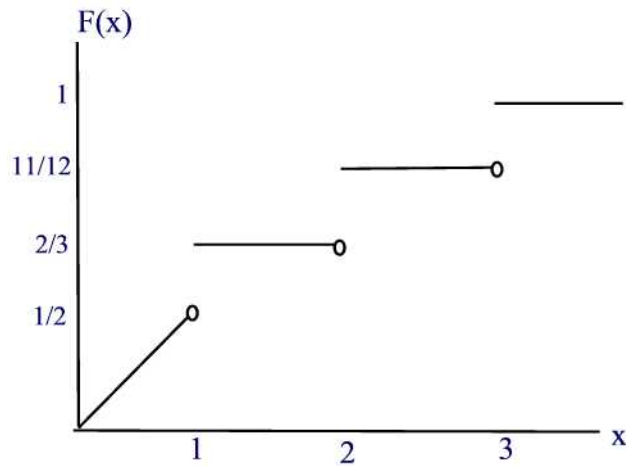


Figure 21.3

Example 21.6

If X has the cdf

$$F(x) = \begin{cases} 0 & \text{for } x < -1 \\ \frac{1}{4} & \text{for } -1 \leq x < 1 \\ \frac{1}{2} & \text{for } 1 \leq x < 3 \\ \frac{3}{4} & \text{for } 3 \leq x < 5 \\ 1 & \text{for } x \geq 5 \end{cases}$$

find (a) $P(X \leq 3)$

(b) $P(X = 3)$

(c) $P(X < 3)$

(d) $P(X \geq 1)$

(e) $P(-0.4 < X < 4)$

(f) $P(-0.4 \leq X < 4)$

(g) $P(-0.4 < X \leq 4)$

(h) $P(-0.4 \leq X \leq 4)$

(i) $P(X = 5)$.

Solution.

(a) $P(X \leq 3) = F(3) = \frac{3}{4}$.

(b) $P(X = 3) = F(3) - F(3^-) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$

(c) $P(X < 3) = F(3^-) = \frac{1}{2}$

(d) $P(X \geq 1) = 1 - F(1^-) = 1 - \frac{1}{4} = \frac{3}{4}$

(e) $P(-0.4 < X < 4) = F(4^-) - F(-0.4) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$

$$(f) P(-0.4 \leq X < 4) = F(4^-) - F(-0.4^-) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$$

$$(g) P(-0.4 < X \leq 4) = F(4) - F(-0.4) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$$

$$(h) P(-0.4 \leq X \leq 4) = F(4) - F(-0.4^-) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$$

$$(i) P(X = 5) = F(5) - F(5^-) = 1 - \frac{3}{4} = \frac{1}{4} \blacksquare$$

Example 21.7

The probability mass function of X , the weekly number of accidents at a certain intersection is given by $p(0) = 0.40$, $p(1) = 0.30$, $p(2) = 0.20$, and $p(3) = 0.10$.

(a) Find the cdf of X .

(b) Find the probability that there will be at least two accidents in any one week.

Solution.

(a) The cdf is given by

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ 0.40 & \text{for } 0 \leq x < 1 \\ 0.70 & \text{for } 1 \leq x < 2 \\ 0.90 & \text{for } 2 \leq x < 3 \\ 1 & \text{for } x \geq 3 \end{cases}$$

$$(b) P(X \geq 2) = 1 - F(2^-) = 1 - 0.70 = 0.30 \blacksquare$$

Problems

Problem 21.1

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.

- Give (explicitly) the probability mass function for X .
- Give (explicitly) the cdf, $F(x)$, for X .
- How much money do you expect to draw from your pocket?

Problem 21.2

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 21.3

Suppose that the cumulative distribution function is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{4} & 0 \leq x < 1 \\ \frac{1}{2} + \frac{x-1}{4} & 1 \leq x < 2 \\ \frac{11}{12} & 2 \leq x < 3 \\ 1 & 3 \leq x \end{cases}$$

- Find $P(X = i)$, $i = 1, 2, 3$.
- Find $P(\frac{1}{2} < X < \frac{3}{2})$.

Problem 21.4

If the cumulative distribution function is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{3} & 0 \leq x < 1 \\ \frac{2}{3} & 1 \leq x < 2 \\ \frac{4}{5} & 2 \leq x < 3 \\ \frac{9}{10} & 3 \leq x < 3.5 \\ 1 & 3.5 \leq x \end{cases}$$

Calculate the probability mass function.

Problem 21.5

Consider a random variable X whose distribution function (cdf) is given by

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.1 & -2 \leq x < 1.1 \\ 0.3 & 1.1 \leq x < 2 \\ 0.6 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

- Give the probability mass function, $p(x)$, of X , explicitly.
- Compute $P(2 < X < 3)$.
- Compute $P(X \geq 3)$.
- Compute $P(X \geq 3 | X \geq 0)$.

Problem 21.6

Consider a random variable X whose probability mass function is given by

$$p(x) = \begin{cases} p & x = -1.9 \\ 0.1 & x = -0.1 \\ 0.3 & x = 20p \\ p & x = 3 \\ 4p & x = 4 \\ 0 & \text{otherwise} \end{cases}$$

- What is p ?
- Find $F(x)$ and sketch its graph.
- What is $F(0)$? What is $F(2)$? What is $F(F(3.1))$?
- What is $P(2X - 3 \leq 4 | X \geq 2.0)$?
- Compute $E(F(X))$.

Problem 21.7

The cdf of X is given by

$$F(x) = \begin{cases} 0 & x < -4 \\ 0.3 & -4 \leq x < 1 \\ 0.7 & 1 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

- Find the probability mass function.
- Find the variance and the standard deviation of X .

Problem 21.8

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.

- Find the probability mass function $p(x)$.
- Compute the cdf $F(x)$ for all numbers x .
- Find the probability that $X < 4$. Is this the same as $F(4)$?

Problem 21.9

A random variable X has cumulative distribution function

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x^2}{4} & 0 \leq x < 1 \\ \frac{1+x}{4} & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

- What is the probability that $X = 0$? What is the probability that $X = 1$? What is the probability that $X = 2$?
- What is the probability that $\frac{1}{2} < X \leq 1$?
- What is the probability that $\frac{1}{2} \leq X < 1$?
- What is the probability that $X > 1.5$?

Hint: You may find it helpful to sketch this function.

Problem 21.10

Let X be a random variable with the following cumulative distribution function:

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

- Find $P(X > \frac{3}{2})$.
- Find $P(\frac{1}{4} < X \leq \frac{3}{4})$.
- Find α .
- Find $P(X = \frac{1}{2})$.
- Sketch the graph of $F(x)$.

Problem 21.11

Suppose the probability function describing the life (in hours) of an ordinary 60 Watt bulb is

$$f(x) = \begin{cases} \frac{1}{1000}e^{-\frac{x}{1000}} & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Let A be the event that the bulb lasts 500 hours or more, and let B be the event that it lasts between 400 and 800 hours.

- (a) What is the probability of A ?
- (b) What is the probability of B ?
- (c) What is $P(A|B)$?

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.

22 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

$$P(X \in B) = \int_B f(x)dx.$$

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

$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

That is, areas under the probability density function represent probabilities as illustrated in Figure 22.1.

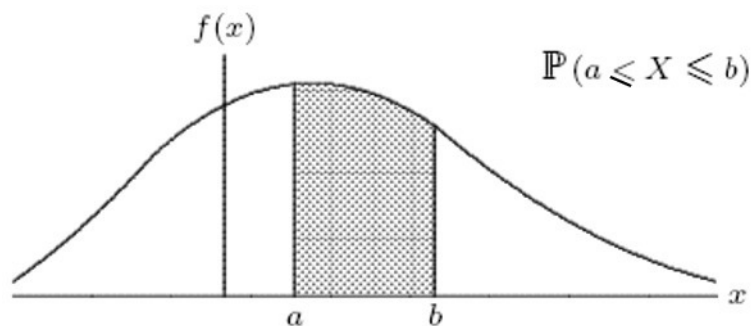


Figure 22.1

Now, if we let $a = b$ in the previous formula we find

$$P(X = a) = \int_a^a f(x)dx = 0.$$

It follows from this result that

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

and

$$P(X \leq a) = P(X < a) \text{ and } P(X \geq a) = P(X > a).$$

The **cumulative distribution function** (abbreviated cdf) $F(t)$ of the random variable X is defined as follows

$$F(t) = P(X \leq t)$$

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

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

Geometrically, $F(t)$ is the area under the graph of f to the left of t .

Example 22.1

If we think of an electron as a particle, the function

$$F(r) = 1 - (2r^2 + 2r + 1)e^{-2r}$$

is the cumulative distribution function of the distance, r , of the electron in a hydrogen atom from the center of the atom. The distance is measured in Bohr radii. (1 Bohr radius = 5.29×10^{-11} m.) Interpret the meaning of $F(1)$.

Solution.

$F(1) = 1 - 5e^{-2} \approx 0.32$. This number says that the electron is within 1 Bohr radius from the center of the atom 32% of the time. ■

Example 22.2

Find the distribution functions corresponding to the following density functions:

$$\begin{aligned} (a) f(x) &= \frac{1}{\pi(1+x^2)}, & -\infty < x < \infty \\ (b) f(x) &= \frac{e^{-x}}{(1+e^{-x})^2}, & -\infty < x < \infty \\ (c) f(x) &= \frac{a-1}{(1+x)^a}, & 0 < x < \infty \\ (d) f(x) &= k\alpha x^{\alpha-1} e^{-kx^\alpha}, & 0 < x < \infty, k > 0, \alpha > 0. \end{aligned}$$

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 $x \geq 0$

$$\begin{aligned} F(x) &= \int_0^x k\alpha y^{\alpha-1} e^{-ky^\alpha} dy \\ &= [-e^{-ky^\alpha}]_0^x \\ &= 1 - ke^{-kx^\alpha} \end{aligned}$$

For $x < 0$ we have $F(x) = 0$ so that

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - ke^{-kx^\alpha} & x \geq 0 \blacksquare \end{cases}$$

Next, we list the properties of the cumulative distribution function $F(t)$ for a continuous random variable X .

Theorem 22.1

The cumulative distribution function of a continuous random variable X satisfies the following properties:

- (a) $0 \leq F(t) \leq 1$.
- (b) $F'(t) = f(t)$.
- (c) $F(t)$ is a non-decreasing function, i.e. if $a < b$ then $F(a) \leq F(b)$.
- (d) $F(t) \rightarrow 0$ as $t \rightarrow -\infty$ and $F(t) \rightarrow 1$ as $t \rightarrow \infty$.
- (e) $P(a < X \leq b) = F(b) - F(a)$.

Proof.

For (a), (c), (d), and (e) refer to Section 21. Part (b) is the result of applying the Second Fundamental Theorem of Calculus to the function $F(t) = \int_{-\infty}^t f(t)dt$ ■

Figure 22.2 illustrates a representative cdf.

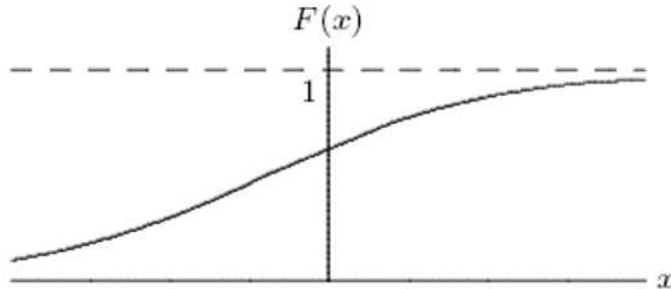


Figure 22.2

Remark 22.1

The intuitive interpretation of the p.d.f. is that for small ϵ

$$P(a \leq X \leq a + \epsilon) = F(a + \epsilon) - F(a) = \int_a^{a+\epsilon} f(t)dt \approx \epsilon f(a)$$

In particular,

$$P\left(a - \frac{\epsilon}{2} \leq X \leq a + \frac{\epsilon}{2}\right) = \epsilon f(a).$$

This means that the probability that X will be contained in an interval of length ϵ around the point a is approximately $\epsilon f(a)$. Thus, $f(a)$ is a measure of how likely it is that the random variable will be near a .

Remark 22.2

By Theorem 22.1 (b) and (d), $\lim_{t \rightarrow -\infty} f(t) = 0 = \lim_{t \rightarrow \infty} f(t)$. This follows from the fact that the graph of $F(t)$ levels off when $t \rightarrow \pm\infty$. That is, $\lim_{t \rightarrow \pm\infty} F'(t) = 0$.

The density function for a continuous random variable, the model for some real-life population of data, will usually be a smooth curve as shown in Figure

22.3.

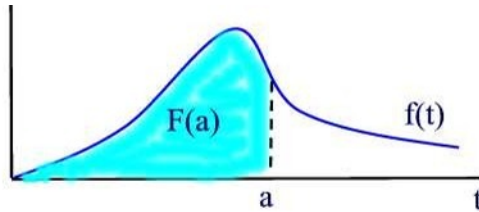


Figure 22.3

Example 22.3

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 $P(-10 \leq X \leq 10)$.

Solution.

$$\begin{aligned} P(-10 \leq X \leq 10) &= \int_{-10}^{10} f(t) dt \\ &= \int_{-10}^0 f(t) dt + \int_0^{10} f(t) dt \\ &= \int_0^{10} e^{-t} dt \\ &= -e^{-t} \Big|_0^{10} = 1 - e^{-10} \blacksquare \end{aligned}$$

Example 22.4

Suppose the income (in tens of thousands of dollars) of people in a community can be approximated by a continuous distribution with density

$$f(x) = \begin{cases} 2x^{-2} & \text{if } x \geq 2 \\ 0 & \text{if } x < 2 \end{cases}$$

Find the probability that a randomly chosen person has an income between \$30,000 and \$50,000.

Solution.

Let X be the income of a randomly chosen person. The probability that a

randomly chosen person has an income between \$30,000 and \$50,000 is

$$P(3 \leq X \leq 5) = \int_3^5 f(x)dx = \int_3^5 2x^{-2}dx = -2x^{-1} \Big|_3^5 = \frac{4}{15} \blacksquare$$

A pdf need not be continuous, as the following example illustrates.

Example 22.5

(a) Determine the value of c so that the following function is a pdf.

$$f(x) = \begin{cases} \frac{15}{64} + \frac{x}{64} & -2 \leq x \leq 0 \\ \frac{3}{8} + cx & 0 < x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

(b) Determine $P(-1 \leq X \leq 1)$.

(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 $P(-1 \leq X \leq 1)$ is calculated as follows.

$$P(-1 \leq X \leq 1) = \int_{-1}^0 \left(\frac{15}{64} + \frac{x}{64} \right) dx + \int_0^1 \left(\frac{3}{8} - \frac{x}{8} \right) dx = \frac{69}{128}$$

(c) For $-2 \leq x \leq 0$ 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 $0 < x \leq 3$

$$F(x) = \int_{-2}^0 \left(\frac{15}{64} + \frac{x}{64} \right) dx + \int_0^x \left(\frac{3}{8} - \frac{t}{8} \right) dt = \frac{7}{16} + \frac{3}{8}x - \frac{1}{16}x^2.$$

Hence the full cdf is

$$F(x) = \begin{cases} 0 & x < -2 \\ \frac{x^2}{128} + \frac{15}{64}x + \frac{7}{16} & -2 \leq x \leq 0 \\ \frac{7}{16} + \frac{3}{8}x - \frac{1}{16}x^2 & 0 < x \leq 3 \\ 1 & x > 3 \end{cases}$$

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 ■

Problems

Problem 22.1

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 22.2

Let X denote the length of time (in minutes) that a certain young lady speaks on the telephone. Assume that the pdf of X is given by

$$f(x) = \begin{cases} \frac{1}{5}e^{-\frac{x}{5}} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- Find the probability that the young lady will talk on the telephone for more than 10 minutes.
- Find the probability that the young lady will talk on the telephone between 5 and 10 minutes.
- Find $F(x)$.

Problem 22.3

A college professor never finishes his lecture before the bell rings to end the period and always finishes his lecture within ten minutes after the bell rings. Suppose that the time X that elapses between the bell and the end of the lecture has 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.

Determine the value of k and find the probability the lecture ends less than 3 minutes after the bell.

Problem 22.4

Let X denote the lifetime (in hours) of a light bulb, and assume that the density function of X is given by

$$f(x) = \begin{cases} 2x & 0 \leq x < \frac{1}{2} \\ \frac{3}{4} & 2 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

On average, what fraction of light bulbs last more than 15 minutes?

Problem 22.5

Define the function $F : \mathbb{R} \rightarrow \mathbb{R}$ 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) Is F the cdf of a continuous random variable? Explain your answer.
 (b) If your answer to part (a) is “yes”, determine the corresponding pdf; if your answer was “no”, then make a modification to F so that it is a cdf, and then compute the corresponding pdf.

Problem 22.6

The amount of time it takes a driver to react to a changing stoplight varies from driver to driver, and is (roughly) described by the continuous probability function of the form:

$$f(x) = \begin{cases} kxe^{-x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant and x is measured in seconds.

- (a) Find the constant k .
 (b) Find the probability that a driver takes more than 1 second to react.

Problem 22.7

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 $P(0 \leq X \leq 1)$.

Problem 22.8

A filling station is supplied with gasoline once a week. If its weekly volume of sales in thousands of gallons is a random variable with pdf

$$f(x) = \begin{cases} 5(1-x)^4 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

What need the capacity of the tank be so that the probability of the supply's being exhausted in a given week is 0.1?

Problem 22.9 ‡

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 22.10 ‡

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 22.11 ‡

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 22.12 ‡

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 22.13 †

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 22.14

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 $P(Y > \frac{1}{2})$.

Problem 22.15

Let X have the density function

$$f(x) = \begin{cases} \frac{3x^2}{\theta^3} & 0 < x < \theta \\ 0 & \text{otherwise} \end{cases}$$

If $P(X > 1) = \frac{7}{8}$, find the value of θ .

Problem 22.16

Insurance claims are for random amounts with a distribution function $F(x)$. Suppose that claims which are no greater than a are no longer paid, but that the underlying distribution of claims remains the same. Write down the distribution function of the claims paid, in terms of $F(x)$. What is the distribution function for the claims not paid?

23 Expectation, Variance and Standard Deviation

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 $x_i = a + i\Delta x, i = 0, 1, \dots, n-1$, be the junction points between the subintervals. Then, the probability of X assuming a value on $[x_i, x_{i+1}]$ is

$$P(x_i \leq X \leq x_{i+1}) = \int_{x_i}^{x_{i+1}} f(x)dx \approx \Delta x f\left(\frac{x_i + x_{i+1}}{2}\right)$$

where we used the midpoint rule to estimate the integral. An estimate of the desired expectation is approximately

$$E(X) \approx \sum_{i=0}^{n-1} \left(\frac{x_i + x_{i+1}}{2}\right) \Delta x f\left(\frac{x_i + x_{i+1}}{2}\right).$$

As n gets ever larger, we see that this converges to the integral

$$E(X) \approx \int_a^b x f(x)dx.$$

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.

From the above discussion it makes sense to 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 23.1

Find $E(X)$ when the density function of X is

$$f(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

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 23.2

The thickness of a conductive coating in micrometers has a density function of $f(x) = 600x^{-2}$ for $100\mu m < x < 120\mu m$. Let X denote the coating thickness.

- Determine the mean and variance of X .
- If the coating costs \$0.50 per micrometer of thickness on each part, what is the average cost of the coating per part?
- Find the probability that the coating thickness exceeds $110\mu m$.

Solution.

(a) Note that $f(x) = 600x^{-2}$ for $100\mu m < x < 120\mu m$ and 0 elsewhere. So, by definition, we have

$$E(X) = \int_{100}^{120} x \cdot 600x^{-2} dx = 600 \ln x \Big|_{100}^{120} \approx 109.39$$

and

$$\sigma^2 = E(X^2) - (E(X))^2 = \int_{100}^{120} x^2 \cdot 600x^{-2} dx - 109.39^2 \approx 33.19.$$

- The average cost per part is $\$0.50 \times (109.39) = \54.70 .
- The desired probability is

$$P(X > 110) = \int_{110}^{120} 600x^{-2} dx = \frac{5}{11} \blacksquare$$

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 23.1

Let Y be a continuous random variable with probability density function f . Then

$$E(Y) = \int_0^{\infty} P(Y > y) dy - \int_{-\infty}^0 P(Y < y) dy.$$

Proof.

From the definition of $E(Y)$ we have

$$\begin{aligned} E(Y) &= \int_0^\infty xf(x)dx + \int_{-\infty}^0 xf(x)dx \\ &= \int_0^\infty \int_{y=0}^{y=x} dyf(x)dx - \int_{-\infty}^0 \int_{y=x}^{y=0} dyf(x)dx \end{aligned}$$

Interchanging the order of integration as shown in Figure 23.1 we can write

$$\int_0^\infty \int_{y=0}^{y=x} dyf(x)dx = \int_0^\infty \int_y^\infty f(x)dx dy$$

and

$$\int_{-\infty}^0 \int_{y=x}^{y=0} dyf(x)dx = \int_{-\infty}^0 \int_{-\infty}^y f(x)dx dy.$$

The result follows by putting the last two equations together and recalling that

$$\int_y^\infty f(x)dx = P(Y > y) \text{ and } \int_{-\infty}^y f(x)dx = P(Y < y) \blacksquare$$

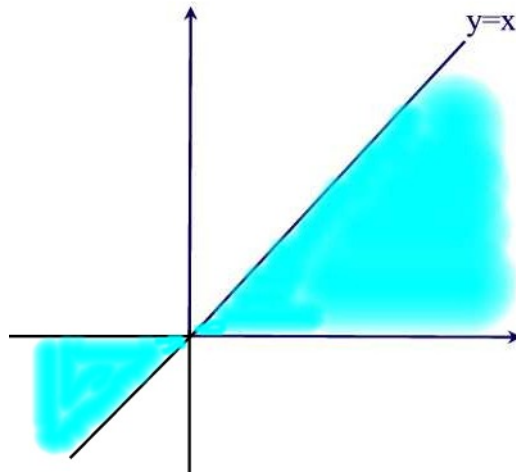


Figure 23.1

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 23.2

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

$$P[g(X) > y] = \int_{B_y} f(x)dx = \int_{\{x:g(x)>y\}} f(x)dx$$

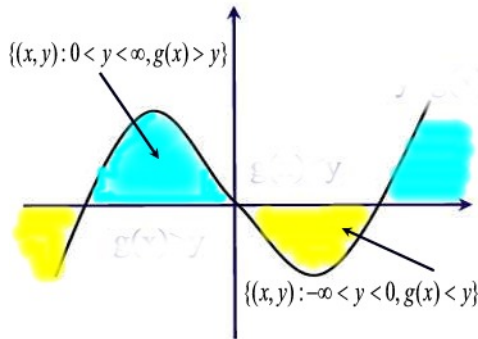
Thus,

$$E(g(X)) = \int_0^{\infty} \left[\int_{\{x:g(x)>y\}} f(x)dx \right] dy - \int_{-\infty}^0 \left[\int_{\{x:g(x)<y\}} f(x)dx \right] dy$$

Now we can interchange the order of integration to obtain

$$\begin{aligned} E(g(X)) &= \int_{\{x:g(x)>0\}} \int_0^{g(x)} f(x)dydx - \int_{\{x:g(x)<0\}} \int_{g(x)}^0 f(x)dydx \\ &= \int_{\{x:g(x)>0\}} g(x)f(x)dx + \int_{\{x:g(x)<0\}} g(x)f(x)dx = \int_{-\infty}^{\infty} g(x)f(x)dx \end{aligned}$$

The following graph of $g(x)$ might help understanding the process of interchanging the order of integration that we used in the proof above ■



Example 23.3

Assume the time T , measured in years, until failure of an appliance has density

$$f(t) = \frac{1}{10}e^{-\frac{t}{10}}, t \geq 0.$$

The manufacturer offers a warranty that pays an amount 100 if the appliance fails during the first year, an amount 50 if it fails during the second or third year, and nothing if it fails after the third year. Determine the expected warranty payment.

Solution.

Let X denote the warranty payment by the manufacturer. We have

$$X = \begin{cases} 100 & 0 < T \leq 1 \\ 50 & 1 < T \leq 3 \\ 0 & T > 3 \end{cases}$$

Thus,

$$\begin{aligned} E(X) &= \int_0^1 100 \frac{1}{10} e^{-\frac{t}{10}} dt + \int_1^3 50 \frac{1}{10} e^{-\frac{t}{10}} dt \\ &= 100(1 - e^{-\frac{1}{10}}) + 50(e^{-\frac{1}{10}} - e^{-\frac{3}{10}}) \\ &= 100 - 50e^{-\frac{1}{10}} - 50e^{-\frac{3}{10}} \blacksquare \end{aligned}$$

Example 23.4 ‡

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 the previous theorem we have

Corollary 23.1

For any constants a and b

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

Proof.

Let $g(x) = ax + b$ in the previous theorem 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 23.5 †

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

$$\begin{aligned} F_Y(y) &= P[(X_1 \leq y) \cap (X_2 \leq y) \cap (X_3 \leq y)] \\ &= P(X_1 \leq y)P(X_2 \leq y)P(X_3 \leq y) \\ &= \left(1 - \frac{1}{y^3}\right)^3, \quad y > 1 \end{aligned}$$

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,

$$\begin{aligned} E(Y) &= \int_1^\infty \left(\frac{9}{y^3}\right) \left(1 - \frac{1}{y^3}\right)^2 dy = \int_1^\infty \left(\frac{9}{y^3}\right) \left(1 - \frac{2}{y^3} + \frac{1}{y^6}\right) dy \\ &= \int_1^\infty \left(\frac{9}{y^3} - \frac{18}{y^6} + \frac{9}{y^9}\right) dy = \left[-\frac{9}{2y^2} + \frac{18}{5y^5} - \frac{9}{y^8}\right]_1^\infty \\ &= 9 \left[\frac{1}{2} - \frac{2}{5} + \frac{1}{8}\right] \approx 2.025 \text{ (in thousands) } \blacksquare \end{aligned}$$

Example 23.6 ‡

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,

$$\text{Var}(X) = E[(X - E(X))^2].$$

Theorem 23.3

(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 23.2 we have

$$\begin{aligned} \text{Var}(X) &= \int_{-\infty}^{\infty} (x - E(X))^2 f(x) dx \\ &= \int_{-\infty}^{\infty} (x^2 - 2xE(X) + (E(X))^2) f(x) dx \\ &= \int_{-\infty}^{\infty} x^2 f(x) dx - 2E(X) \int_{-\infty}^{\infty} x f(x) dx + (E(X))^2 \int_{-\infty}^{\infty} f(x) dx \\ &= E(X^2) - (E(X))^2 \end{aligned}$$

(b) We have

$$\text{Var}(aX + b) = E[(aX + b - E(aX + b))^2] = E[a^2(X - E(X))^2] = a^2\text{Var}(X) \blacksquare$$

Example 23.7

Let X be a random variable with probability density function

$$f(x) = \begin{cases} 2 - 4|x| & -\frac{1}{2} < x < \frac{1}{2}, \\ 0 & \text{elsewhere} \end{cases}$$

- (a) Find the variance of X .
 (b) Find the c.d.f. $F(x)$ of X .

Solution.

(a) By the symmetry of the distribution about 0, $E(X) = 0$. Thus,

$$\begin{aligned} \text{Var}(X) = E(X^2) &= \int_{-\frac{1}{2}}^0 x^2(2 + 4x)dx + \int_0^{\frac{1}{2}} x^2(2 - 4x)dx \\ &= 2 \int_0^{\frac{1}{2}} (2x^2 - 8x^3)dx = \frac{1}{24} \end{aligned}$$

(b) Since the range of f is the interval $(-\frac{1}{2}, \frac{1}{2})$, we have $F(x) = 0$ for $x \leq -\frac{1}{2}$ and $F(x) = 1$ for $x \geq \frac{1}{2}$. Thus it remains to consider the case when $-\frac{1}{2} < x < \frac{1}{2}$. For $-\frac{1}{2} < x \leq 0$,

$$F(x) = \int_{-\frac{1}{2}}^x (2 + 4t)dt = 2x^2 + 2x + \frac{1}{2}$$

For $0 \leq x < \frac{1}{2}$, 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

$$F(x) = \begin{cases} 0 & x < -\frac{1}{2} \\ 2x^2 + 2x + \frac{1}{2} & -\frac{1}{2} \leq x < 0 \\ -2x^2 + 2x + \frac{1}{2} & 0 \leq x < \frac{1}{2} \\ 1 & x \geq \frac{1}{2} \blacksquare \end{cases}$$

Example 23.8

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

$$P(X < 1) = P(X \leq 1) = \int_0^1 4xe^{-2x} dx = \int_0^2 te^{-t} dt = -(t+1)e^{-t} \Big|_0^2 = 1 - 3e^{-2} \blacksquare$$

As in the case of discrete random variable, it is easy to establish the formula

$$\text{Var}(aX) = a^2 \text{Var}(X)$$

Example 23.9

Suppose that the cost of maintaining a car is given by a random variable, X , with mean 200 and variance 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 $\text{Var}(1.2X) = 1.2^2\text{Var}(X) = (1.44)(260) = 374$. ■

Next, we define the **standard deviation** X to be the square root of the variance.

Example 23.10

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

$$E(X) = \int_{x_0}^{\infty} x f(x) dx = \int_{x_0}^{\infty} \frac{\alpha x_0^\alpha}{x^\alpha} dx = \frac{\alpha}{1-\alpha} \left(\frac{x_0^\alpha}{x^{\alpha-1}}\right)\Big|_{x_0}^{\infty} = \frac{\alpha x_0}{\alpha-1}$$

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,

$$\text{Var}(X) = \frac{\alpha x_0^2}{\alpha-2} - \frac{\alpha^2 x_0^2}{(\alpha-1)^2} = \frac{\alpha x_0^2}{(\alpha-2)(\alpha-1)^2} \quad \blacksquare$$

Percentiles and Interquartile Range

We define the **100pth percentile** of a population to be the quantity that

separates the lower $100p\%$ of a population from the upper $100(1 - p)\%$. For a random variable X , the $100p^{\text{th}}$ percentile is the smallest number x such that

$$F(x) \geq p. \quad (23.3)$$

The 50^{th} percentile is called the **median** of X . In the case of a continuous random variable, (23.3) holds with equality; in the case of a discrete random variable, equality may not hold.

The difference between the 75^{th} percentile and 25^{th} percentile is called the **interquartile range**.

Example 23.11

Suppose the random variable X has pmf

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

Find the median and the 70^{th} percentile.

Solution.

We have

$$F(0) = \frac{1}{3} \approx 0.33$$

$$F(1) = \frac{1}{3} + \frac{2}{9} \approx 0.56$$

$$F(2) = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} \approx 0.70$$

Thus, the median is 1 and the 70^{th} percentile is 2 ■

Example 23.12

Suppose the random variable X has pdf

$$f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find the $100p^{\text{th}}$ percentile.

Solution.

The cdf of X is given by

$$F(x) = \int_0^x e^{-t} dt = 1 - e^{-x}$$

Thus, the 100pth percentile is the solution to the equation

$$1 - e^{-x} = p.$$

For example, for the median, $p = 0.5$ so that

$$1 - e^{-x} = 0.5$$

Solving for x we find $x \approx 0.693$ ■

Problems

Problem 23.1

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 $P(0 \leq x \leq 0.5)$.
- (d) Find $E(X)$.

Problem 23.2

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 also that you are told that $E(X) = \frac{3}{5}$.

- (a) Find a and b .
- (b) Determine the cdf, $F(x)$, explicitly.

Problem 23.3

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 23.4

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 23.5

The lifetime (in years) of a certain brand of light bulb is a random variable (call it X), with probability density function

$$f(x) = \begin{cases} 4(1+x)^{-5} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Find the mean and the standard deviation.
 (b) What is the probability that a randomly chosen lightbulb lasts less than a year?

Problem 23.6

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 23.7

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 23.8

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 23.9 ‡

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 23.10 ‡

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 23.11 ‡

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 $0 < x < \infty$.

Determine the company's expected monthly claims.

Problem 23.12 ‡

An insurer's annual weather-related loss, X , is a random variable with 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 25th and 75th percentiles of X .

Problem 23.13 ‡

A random variable X has the cumulative distribution function

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

Calculate the variance of X .

Problem 23.14 ‡

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?

Problem 23.15 ‡

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 23.16

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 23.17

Let X be a random 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 23.18

Let $f(x)$ be the density function of a random variable X . We define the **mode** of X to be the number that maximizes $f(x)$. Let X be a continuous random variable with density function

$$f(x) = \begin{cases} \lambda^{\frac{1}{3}} x(4-x) & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

Find the mode of X .

Problem 23.19

A system made up of 7 components with independent, identically distributed lifetimes will operate until any of 1 of the system's components fails. If the lifetime X of each component has density function

$$f(x) = \begin{cases} \lambda \frac{3}{x^4} & x > 1 \\ 0 & \text{otherwise} \end{cases}$$

What is the expected lifetime until failure of the system?

Problem 23.20

Let X have the density function

$$f(x) = \begin{cases} \lambda \frac{2x}{k^2} & 0 \leq x \leq k \\ 0 & \text{otherwise} \end{cases}$$

For what value of k is the variance of X equal to 2?

Problem 23.21

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 23.22

A game is played where a player generates 4 independent Uniform(0,100) random variables and wins the maximum of the 4 numbers.

- (a) Give the density function of a player's winnings.
- (b) What is the player expected winnings?

24 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 $F(x) = \int_{-\infty}^x f(t)dt$ then 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 24.1 presents a graph of $f(x)$ and $F(x)$.

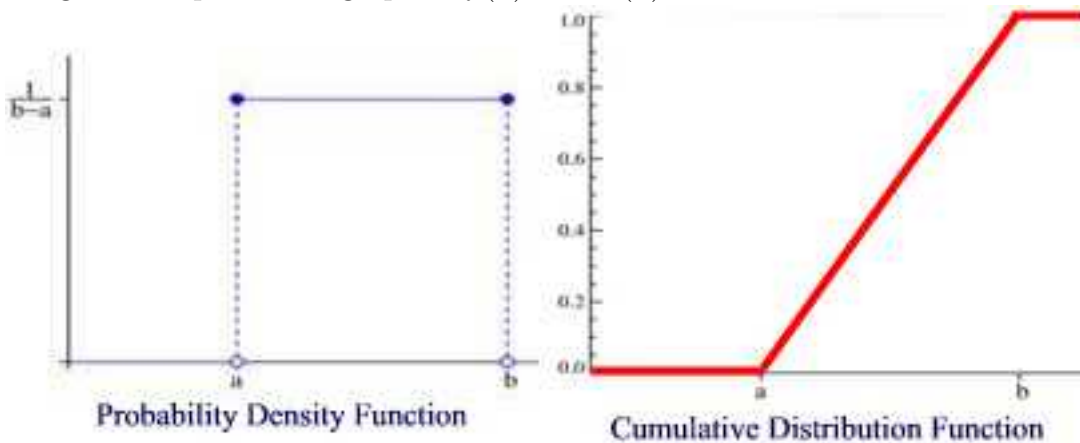


Figure 24.1

If $a = 0$ and $b = 1$ then X is called the **standard uniform** random variable.

Remark 24.1

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 24.1

You are the production manager of a soft drink bottling company. You believe that when a machine is set to dispense 12 oz., it really dispenses 11.5 to 12.5 oz. inclusive. Suppose the amount dispensed has a uniform distribution. What is the probability that less than 11.8 oz. is dispensed?

Solution.

Since $f(x) = \frac{1}{12.5-11.5} = 1$ then

$$P(11.5 \leq X \leq 11.8) = \text{area of rectangle of base } 0.3 \text{ and height } 1 = 0.3 \blacksquare$$

Example 24.2

Suppose that X has a uniform distribution on the interval $(0, a)$, where $a > 0$. Find $P(X > X^2)$.

Solution.

If $a \leq 1$ then $P(X > X^2) = \int_0^a \frac{1}{a} dx = 1$. If $a > 1$ then $P(X > X^2) = \int_0^1 \frac{1}{a} dx = \frac{1}{a}$. Thus, $P(X > X^2) = \min\{1, \frac{1}{a}\} \blacksquare$

The expected value of X is

$$\begin{aligned} E(X) &= \int_a^b x f(x) \\ &= \int_a^b \frac{x}{b-a} dx \\ &= \frac{x^2}{2(b-a)} \Big|_a^b \\ &= \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2} \end{aligned}$$

and so the expected value of a uniform random variable is halfway between a and b . Because

$$\begin{aligned} E(X^2) &= \int_a^b \frac{x^2}{b-a} dx \\ &= \frac{x^3}{3(b-a)} \Big|_a^b \\ &= \frac{b^3 - a^3}{3(b-a)} \\ &= \frac{a^2 + b^2 + ab}{3} \end{aligned}$$

then

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{a^2 + b^2 + ab}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}.$$

Example 24.3

You arrive at a bus stop 10:00 am, knowing that the bus will arrive at some time uniformly distributed between 10:00 and 10:30 am. Let X be your wait time for the bus. Then X is uniformly distributed in $(0, 30)$. Find $E(X)$ and $\text{Var}(X)$.

Solution.

We have $a = 0$ and $b = 30$. Thus, $E(X) = \frac{a+b}{2} = 15$ and $\text{Var}(X) = \frac{(b-a)^2}{12} = \frac{30^2}{12} = 75$ ■

Problems

Problem 24.1

The total time to process a loan application is uniformly distributed between 3 and 7 days.

- Let X denote the time to process a loan application. Give the mathematical expression for the probability density function.
- What is the probability that a loan application will be processed in fewer than 3 days ?
- What is the probability that a loan application will be processed in 5 days or less ?

Problem 24.2

Customers at Rositas Cocina are charged for the amount of salad they take. Sampling suggests that the amount of salad taken is uniformly distributed between 5 ounces and 15 ounces. Let X = salad plate filling weight

- Find the probability density function of X .
- What is the probability that a customer will take between 12 and 15 ounces of salad?
- Find $E(X)$ and $\text{Var}(X)$.

Problem 24.3

Suppose that X has a uniform distribution over the interval $(0, 1)$. Find

- $F(x)$;
- Show that $P(a \leq X \leq a + b)$ for $a, b \geq 0$, $a + b \leq 1$ depends only on b .

Problem 24.4

Let X be uniform on $(0,1)$. Compute $E(X^n)$.

Problem 24.5

Let X be a uniform random variable on the interval $(1,2)$ and let $Y = \frac{1}{X}$. Find $E[Y]$.

Problem 24.6

You arrive at a bus stop at 10:00 am, knowing that the bus will arrive at some time uniformly distributed between 10:00 and 10:30. What is the probability that you will have to wait longer than 10 minutes?

Problem 24.7 ‡

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 24.8 ‡

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 24.9 ‡

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 24.10

Let X be a random variable distributed uniformly over the interval $[-1, 1]$.

- (a) Compute $E(e^{-X})$.
- (b) Compute $\text{Var}(e^{-X})$.

Problem 24.11

Let X be a random variable with a continuous uniform distribution on the interval $(1, a)$, $a > 1$. If $E(X) = 6\text{Var}(X)$, what is the value of a ?

Problem 24.12

Let X be a random variable with a continuous uniform distribution on the interval $(0, 10)$. What is $P(X + \frac{10}{X} > 7)$?

Problem 24.13

Suppose that X has a uniform distribution on the interval $(0, a)$, $a > 0$. Find $P(X > X^2)$.

25 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 25.1).

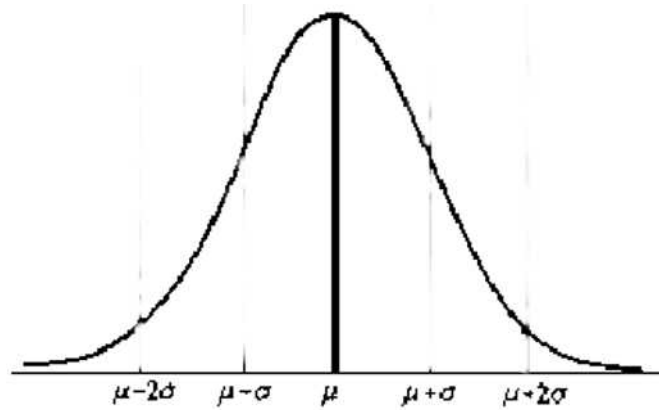


Figure 25.1

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 μ —hence the experiment outcome being modeled should be equally 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 see later, 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

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dx dy \\ &= \int_0^{\infty} \int_0^{2\pi} e^{-\frac{r^2}{2}} r dr d\theta \\ &= 2\pi \int_0^{\infty} r e^{-\frac{r^2}{2}} dr = 2\pi \end{aligned}$$

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 $dydx = r dr d\theta$.

Example 25.1

The width of a bolt of fabric is normally distributed with mean 950mm and standard deviation 10 mm. What is the probability that a randomly chosen bolt has a width of between 947 and 958 mm?

Solution.

Let X be the width (in mm) of a randomly chosen bolt. Then X is normally distributed with mean 950 mm and variance 100 mm. Thus,

$$P(947 \leq X \leq 950) = \frac{1}{10\sqrt{2\pi}} \int_{947}^{950} e^{-\frac{(x-950)^2}{200}} dx \approx 0.4 \blacksquare$$

Theorem 25.1

If X is a normal distribution with parameters (μ, σ^2) then $Y = aX + b$ is a normal distribution with parameters $(a\mu + b, a^2\sigma^2)$.

Proof.

We prove the result when $a > 0$. The proof is similar for $a < 0$. Let F_Y denote the cdf of Y . Then

$$\begin{aligned} F_Y(x) &= P(Y \leq x) = P(aX + b \leq x) \\ &= P\left(X \leq \frac{x-b}{a}\right) = F_X\left(\frac{x-b}{a}\right) \end{aligned}$$

Differentiating both sides to obtain

$$\begin{aligned} f_Y(x) &= \frac{1}{a} f_X\left(\frac{x-b}{a}\right) \\ &= \frac{1}{\sqrt{2\pi}a\sigma} \exp\left[-\left(\frac{x-b}{a} - \mu\right)^2 / (2\sigma^2)\right] \\ &= \frac{1}{\sqrt{2\pi}a\sigma} \exp\left[-(x - (a\mu + b))^2 / 2(a\sigma)^2\right] \end{aligned}$$

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 25.2

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 $dv = x e^{-\frac{x^2}{2}}$ we find

$$\text{Var}(Z) = \frac{1}{\sqrt{2\pi}} \left[-x e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1.$$

Thus,

$$\text{Var}(X) = \text{Var}(\sigma Z + \mu) = \sigma^2 \text{Var}(Z) = \sigma^2 \quad \blacksquare$$

Figure 25.2 shows different normal curves with the same μ and different σ .

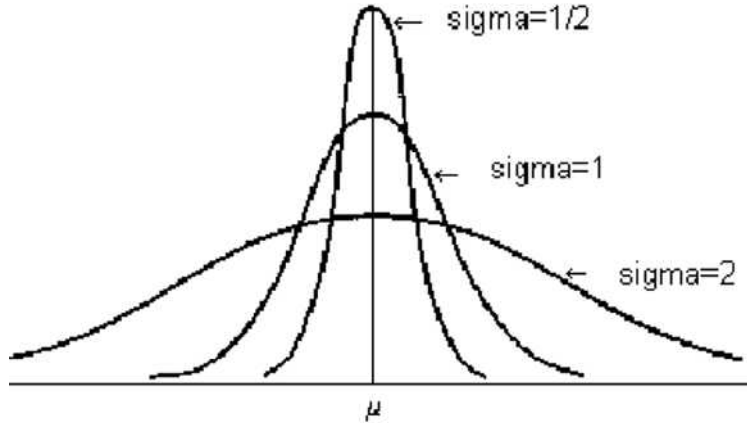


Figure 25.2

Example 25.2

A college has an enrollment of 3264 female students. Records show that the mean height of these students is 64.4 inches and the standard deviation is 2.4 inches. Since the shape of the relative histogram of sample college students approximately normally distributed, we assume the total population distribution of the height X of all the female college students follows the normal distribution with the same mean and the standard deviation. Find $P(66 \leq X \leq 68)$.

Solution.

If you want to find out the percentage of students whose heights are between 66 and 68 inches, you have to evaluate the area under the normal curve from 66 to 68 as shown in Figure 25.3.

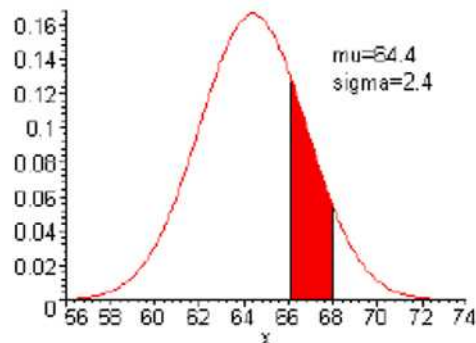


Figure 25.3

Thus,

$$P(66 \leq X \leq 68) = \frac{1}{\sqrt{2\pi}(2.4)} \int_{66}^{68} e^{-\frac{(x-64.4)^2}{2(2.4)^2}} dx \approx 0.1846 \blacksquare$$

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 $f_Z(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ then $f_Z(x)$ 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,

$$\Phi(x) = 1 - \Phi(-x), \quad -\infty < x < \infty. \quad (25.4)$$

This implies that

$$P(Z \leq -x) = P(Z > x).$$

Now, $\Phi(x)$ is the area under the standard curve to the left of x . The values of $\Phi(x)$ for $x \geq 0$ are given in Table 25.1 below. Equation 25.4 is used for $x < 0$.

Example 25.3

On May 5, in a certain city, temperatures have been found to be normally distributed with mean $\mu = 24^\circ\text{C}$ and variance $\sigma^2 = 9$. The record temperature on that day is 27°C .

- What is the probability that the record of 27°C will be broken next May 5?
- What is the probability that the record of 27°C will be broken at least 3 times during the next 5 years on May 5? (Assume that the temperatures during the next 5 years on May 5 are independent.)
- How high must the temperature be to place it among the top 5% of all temperatures recorded on May 5?

Solution.

(a) Let X be the temperature on May 5. Then X has a normal distribution with $\mu = 24$ and $\sigma = 3$. The desired probability is given by

$$\begin{aligned} P(X > 27) &= P\left(\frac{X - 24}{3} > \frac{27 - 24}{3}\right) = P(Z > 1) \\ &= 1 - P(Z \leq 1) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587 \end{aligned}$$

(b) Let Y be the number of times with broken records during the next 5 years on May 5. Then, Y has a binomial distribution with $n = 5$ and $p = 0.1587$. So, the desired probability is

$$\begin{aligned} P(Y \geq 3) &= P(Y = 3) + P(Y = 4) + P(Y = 5) \\ &= C(5, 3)(0.1587)^3(0.8413)^2 + C(5, 4)(0.1587)^4(0.8413)^1 \\ &\quad + C(5, 5)(0.1587)^5(0.8413)^0 \\ &\approx 0.03106 \end{aligned}$$

(c) Let x be the desired temperature. We must have $P(X > x) = 0.05$ or equivalently $P(X \leq x) = 0.95$. Note that

$$P(X \leq x) = P\left(\frac{X - 24}{3} < \frac{x - 24}{3}\right) = P\left(Z < \frac{x - 24}{3}\right) = 0.95$$

From the normal Table below we find $P(Z \leq 1.65) = 0.95$. Thus, we set $\frac{x-24}{3} = 1.65$ and solve for x we find $x = 28.95^\circ\text{C}$ ■

Next, we point out that probabilities involving normal random variables are reduced to the ones involving standard normal variable. For example

$$P(X \leq a) = P\left(\frac{X - \mu}{\sigma} \leq \frac{a - \mu}{\sigma}\right) = \Phi\left(\frac{a - \mu}{\sigma}\right).$$

Example 25.4

Let X be a normal random variable with parameters μ and σ^2 . Find

- (a) $P(\mu - \sigma \leq X \leq \mu + \sigma)$.
- (b) $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma)$.
- (c) $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma)$.

Solution.

(a) We have

$$\begin{aligned} P(\mu - \sigma \leq X \leq \mu + \sigma) &= P(-1 \leq Z \leq 1) \\ &= \Phi(1) - \Phi(-1) \\ &= 2(0.8413) - 1 = 0.6826. \end{aligned}$$

Thus, 68.26% of all possible observations lie within one standard deviation to either side of the mean,

(b) We have

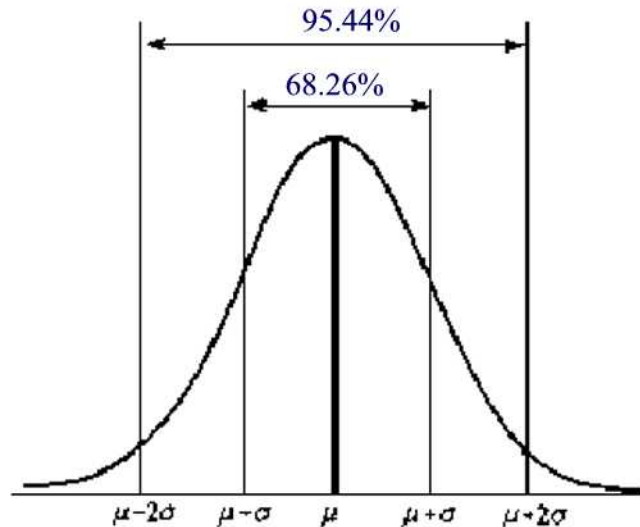
$$\begin{aligned} P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) &= P(-2 \leq Z \leq 2) \\ &= \Phi(2) - \Phi(-2) \\ &= 2(0.9772) - 1 = 0.9544. \end{aligned}$$

Thus, 95.44% of all possible observations lie within two standard deviations to either side of the mean.

(c) We have

$$\begin{aligned} P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) &= P(-3 \leq Z \leq 3) \\ &= \Phi(3) - \Phi(-3) \\ &= 2(0.9987) - 1 = 0.9974. \end{aligned}$$

Thus, 99.74% of all possible observations lie within three standard deviations to either side of the mean ■



The Normal Approximation to the Binomial Distribution

Historically, the normal distribution was discovered by DeMoivre as an approximation to the binomial distribution. The result is the so-called *De Moivre-Laplace theorem*.

Theorem 25.3

Let S_n denote the number of successes that occur with n independent Bernoulli

trials, each with probability p of success. Then, for $a < b$,

$$\lim_{n \rightarrow \infty} P \left[a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b \right] = \Phi(b) - \Phi(a).$$

Proof.

This result is a special case of the central limit theorem, which will be discussed in Section 40. Consequently, we will defer the proof of this result until then ■

Remark 25.1

The normal approximation to the binomial distribution is usually quite good when $np \geq 10$ and $n(1-p) \geq 10$.

Remark 25.2

Suppose we are approximating a binomial random variable with a normal random variable. Say we want to find $P(X \leq 10)$. Figure 25.4 zooms in on the left portion of the distributions and shows $P(X \leq 10)$. Since each histogram bar is centered at an integer, the shaded area actually goes to 10.5. If we were to approximate $P(X \leq 10)$ with the normal cdf evaluated at 10, we would effectively be missing half of the bar centered at 10. In practice, then, we apply a *continuity correction*, by evaluating the normal cdf at 10.5.

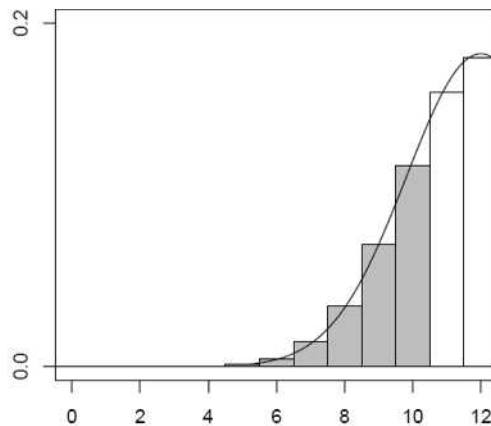


Figure 25.4

Example 25.5

A process yields 10% defective items. If 100 items are randomly selected from the process, 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 $np = 10 \geq 10$ and $n(1 - p) = 90 \geq 10$ we can use the normal approximation to the binomial with $\mu = np = 10$ and $\sigma^2 = np(1 - p) = 9$. We want $P(X > 13)$. Using continuity correction we find

$$\begin{aligned} P(X > 13) &= P(X \geq 14) \\ &= P\left(\frac{X - 10}{\sqrt{9}} \geq \frac{13.5 - 10}{\sqrt{9}}\right) \\ &\approx 1 - \Phi(1.17) = 1 - 0.8790 = 0.121 \blacksquare \end{aligned}$$

Example 25.6

In the United States, $\frac{1}{6}$ of the people are lefthanded. In a small town (a random sample) of 612 persons, 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 $np = 102 \geq 10$ and $n(1 - p) = 510 \geq 10$ 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

$$\begin{aligned} P(90 < X < 150) &= P(91 \leq X \leq 149) = \\ &= P\left(\frac{90.5 - 102}{\sqrt{85}} \leq \frac{X - 102}{\sqrt{85}} \leq \frac{149.5 - 102}{\sqrt{85}}\right) \\ &= P(-1.25 \leq Z \leq 5.15) \approx 0.8943 \end{aligned}$$

The last number is found by using TI83 Plus \blacksquare

Example 25.7

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 or more 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

$$p = \frac{C(4, 2)C(48, 11)}{C(52, 13)} + \frac{C(4, 3)C(48, 10)}{C(52, 13)} + \frac{C(4, 4)C(48, 9)}{C(52, 13)} \approx 0.2573$$

Since $np \approx 23.573 \geq 10$ and $n(1-p) = 66.843 \geq 10$, X can be approximated by a normal random variable with $\mu = 23.573$ and $\sigma = \sqrt{np(1-p)} \approx 4.1473$. Thus,

$$\begin{aligned} P(X > 50) &= 1 - P(X \leq 50) = 1 - \Phi\left(\frac{50.5 - 23.573}{4.1473}\right) \\ &\approx 1 - \Phi(6.5) \blacksquare \end{aligned}$$

Problems

Problem 25.1

Scores for a particular standardized test are Normally distributed with a mean of 80 and a standard deviation of 14. Find the probability that a randomly chosen score is

- (a) no greater than 70
- (b) at least 95
- (c) between 70 and 95.
- (d) A student was told that her percentile score on this exam is 72%. Approximately what is her raw score?

Problem 25.2

Suppose that egg shell thickness is normally distributed with a mean of 0.381 mm and a standard deviation of 0.031 mm.

- (a) Find the proportion of eggs with shell thickness more than 0.36 mm.
- (b) Find the proportion of eggs with shell thickness within 0.05 mm of the mean.
- (c) Find the proportion of eggs with shell thickness more than 0.07 mm from the mean.

Problem 25.3

Assume the time required for a certain distance runner to run a mile follows a normal distribution with mean 4 minutes and variance 4 seconds.

- (a) What is the probability that this athlete will run the mile in less than 4 minutes?
- (b) What is the probability that this athlete will run the mile in between 3 min55sec and 4min5sec?

Problem 25.4

You work in Quality Control for GE. Light bulb life has a normal distribution with $\mu = 2000$ hours and $\sigma^2 = 200$ hours. What's the probability that a bulb will last

- (a) between 2000 and 2400 hours?
- (b) less than 1470 hours?

Problem 25.5

Human intelligence (IQ) has been shown to be normally distributed with mean 100 and standard deviation 15. What fraction of people have IQ greater than 130 ("the gifted cutoff"), given that $\Phi(2) = .9772$?

Problem 25.6

Let X represent the lifetime of a randomly chosen battery. Suppose X is a normal random variable with parameters $(50, 25)$.

- (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 25.7

Suppose that 25% of all licensed drivers do not have insurance. Let X be the number of uninsured drivers in a random sample of 50.

- (a) What is the probability that the number of uninsured drivers is at most 10?
- (b) What is the probability that the number of uninsured drivers is between 5 and 15?

Problem 25.8 †

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 25.9

If for a certain normal random variable X , $P(X < 500) = 0.5$ and $P(X > 650) = 0.0227$, find the standard deviation of X .

Problem 25.10

A computer generates 10,000 random decimal digits $0, 1, 2, \dots, 9$ (each selected with probability $\frac{1}{10}$) and then tabulates the number of occurrences of each digit.

- (a) Using an appropriate approximation, find the (approximate) probability that exactly 1,060 of the 10,000 digits are 0.
- (b) Determine x (as small as possible) such that, with 99.95% probability, the number of digits 0 is at most x .

Problem 25.11

Suppose that X is a normal random variable with parameters $\mu = 5, \sigma^2 = 49$. Using the table of the normal distribution, compute: (a) $P(X > 5.5)$, (b) $P(4 < X < 6.5)$, (c) $P(X < 8)$, (d) $P(|X - 7| \geq 4)$.

Problem 25.12

A company wants to buy boards of length 2 meters and is willing to accept lengths that are off by as much as 0.04 meters. The board manufacturer produces boards of length normally distributed with mean 2.01 meters and standard deviation σ .

If the probability that a board is too long is 0.01, what is σ ?

Problem 25.13

Let X be a normal random variable with mean 1 and variance 4. Find $P(X^2 - 2X \leq 8)$.

Problem 25.14

Scores on a standardized exam are normally distributed with mean 1000 and standard deviation 160.

- (a) What proportion of students score under 850 on the exam?
- (b) They wish to calibrate the exam so that 1400 represents the 98th percentile. What should they set the mean to? (without changing the standard deviation)

Problem 25.15

The daily number of arrivals to a rural emergency room is a Poisson random variable with a mean of 100 people per day. Use the normal approximation to the Poisson distribution to obtain the approximate probability that 112 or more people arrive in a day.

Problem 25.16

A machine is used to automatically fill 355ml pop bottles. The actual amount put into each bottle is a normal random variable with mean 360ml and standard deviation of 4ml.

- (a) What proportion of bottles are filled with less than 355ml of pop?
- b) Suppose that the mean fill can be adjusted. To what value should it be set so that only 2.5% of bottles are filled with less than 355ml?

Problem 25.17

Suppose that your journey time from home to campus is normally distributed with mean equal to 30 minutes and standard deviation equal to 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 class at 2pm?

Problem 25.18

Suppose that a local vote is being held to see if a new manufacturing facility will may be built in the locality. A polling company will survey 200 individuals to measure support for the new facility. If in fact 53% of the population oppose the building of this facility, 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 25.19

Suppose that the current measurements in a strip of wire are assumed to follow a normal distribution with mean of 12 milliamperes and a standard deviation of 9 (milliamperes).

- (a) What is the probability that a measurement will exceed 14 milliamperes?
- (b) What is the probability that a current measurement is between 9 and 16 milliamperes?
- (c) Determine the value for which the probability that a current measurement is below this value is 0.95.

Problem 25.20

Suppose that the current measurements in a strip of wire are assumed to follow a normal distribution with mean of 12 milliamperes and a standard deviation of 9 (milliamperes).

- (a) What is the probability that a measurement will exceed 14 milliamperes?
- (b) What is the probability that a current measurement is between 9 and 16 milliamperes?
- (c) Determine the value for which the probability that a current measurement is below this value is 0.95.

26 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 26.1

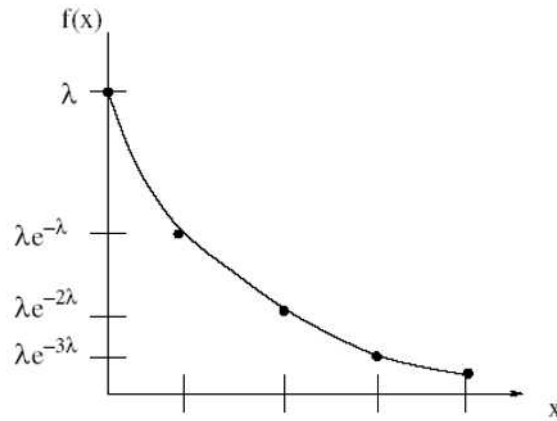


Figure 26.1

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 $dv = \lambda e^{-\lambda x} dx$:

$$\begin{aligned} E(X) &= \int_0^{\infty} x \lambda e^{-\lambda x} dx \\ &= [-x e^{-\lambda x}]_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx \\ &= [-x e^{-\lambda x}]_0^{\infty} + \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_0^{\infty} \\ &= \frac{1}{\lambda} \end{aligned}$$

Furthermore, using integration by parts again, we may also obtain that

$$\begin{aligned} E(X^2) &= \int_0^{\infty} \lambda x^2 e^{-\lambda x} dx = \int_0^{\infty} x^2 d(-e^{-\lambda x}) \\ &= [-x^2 e^{-\lambda x}]_0^{\infty} + 2 \int_0^{\infty} x e^{-\lambda x} dx \\ &= \frac{2}{\lambda^2} \end{aligned}$$

Thus,

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Example 26.1

The time between machine failures at an industrial plant has an exponential distribution with an average of 2 days between failures. Suppose a failure has just occurred at the plant. Find the probability that the next failure won't happen in the next 5 days.

Solution.

Let X denote the time between accidents. The mean time to failure is 2 days. Thus, $\lambda = 0.5$. Now, $P(X > 5) = 1 - P(X \leq 5) = \int_5^{\infty} 0.5e^{-0.5x} dx \approx 0.082085$ ■

Example 26.2

The mileage (in thousands of miles) which car owners get with a certain kind of radial tire is a random variable having an exponential distribution with mean 40. Find the probability that one of these tires will last at most 30 thousand miles.

Solution.

Let X denote the mileage (in thousands of miles) of one these tires. Then X is an exponential distribution with parameter $\lambda = \frac{1}{40}$. Thus,

$$\begin{aligned} P(X \leq 30) &= \int_0^{30} \frac{1}{40} e^{-\frac{x}{40}} dx \\ &= -e^{-\frac{x}{40}} \Big|_0^{30} = 1 - e^{-\frac{3}{4}} \approx 0.5276 \quad \blacksquare \end{aligned}$$

The cumulative distribution function is given by

$$F(x) = P(X \leq x) = \int_0^x \lambda e^{-\lambda u} du = -e^{-\lambda u} \Big|_0^x = 1 - e^{-\lambda x}, \quad x \geq 0$$

Example 26.3

Suppose that the length of a phone call in minutes is an exponential random variable with mean 10 minutes. If someone arrives immediately ahead of you at a public telephone booth, 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 at the phone booth. Then X is an exponential random variable with parameter $\lambda = 0.1$. (a) We have $P(X > 10) = 1 - F(10) = 1 - (1 - e^{-1}) = e^{-1} \approx 0.3679$.

(b) We have $P(10 \leq X \leq 20) = F(20) - F(10) = e^{-1} - e^{-2} \approx 0.2325$ ■

The most important property of the exponential distribution is known as the **memoryless** property:

$$P(X > s + t | X > s) = P(X > t), \quad s, t \geq 0.$$

That 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

$$P(X > t) = \int_t^{\infty} \lambda e^{-\lambda x} dx = -e^{-\lambda x} \Big|_t^{\infty} = e^{-\lambda t}.$$

It follows that

$$\begin{aligned} P(X > s + t | X > s) &= \frac{P(X > s + t \text{ and } X > s)}{P(X > s)} \\ &= \frac{P(X > s + t)}{P(X > s)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} = P(X > t) \end{aligned}$$

Example 26.4

Let X be the time (in hours) required to repair a computer system. We

assume that X has an exponential distribution with parameter $\lambda = \frac{1}{4}$. Find

(a) the cumulative distribution function of X .

(b) $P(X > 4)$.

(c) $P(X > 10|X > 8)$.

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) $P(X > 4) = 1 - P(X \leq 4) = 1 - F(4) = 1 - (1 - e^{-\frac{4}{4}}) = e^{-1} \approx 0.368$.

(c) By the memoryless property, we find

$$\begin{aligned} P(X > 10|X > 8) &= P(X > 8 + 2|X > 8) = P(X > 2) \\ &= 1 - P(X \leq 2) = 1 - F(2) \\ &= 1 - (1 - e^{-\frac{1}{2}}) = e^{-\frac{1}{2}} \approx 0.6065 \blacksquare \end{aligned}$$

Example 26.5

The distance between major cracks in a highway follows an exponential distribution with a mean of 5 miles.

(a) What is the probability that there are no major cracks in a 20-mile stretch of the highway?

(b) What is the probability that the first major crack occurs between 15 and 20 miles of the start of inspection?

(c) Given that there are no cracks in the first 5 miles inspected, what is the probability that there are no major cracks in the next 15 miles?

Solution.

Let X denote the distance between major cracks. Then, X is an exponential random variable with $\mu = \frac{1}{E(X)} = \frac{1}{5} = 0.2$ cracks/mile.

(a)

$$P(X > 20) = \int_{20}^{\infty} 0.2e^{-0.2x} dx = -e^{-0.2x} \Big|_{20}^{\infty} = e^{-4} \approx 0.01831.$$

(b)

$$P(15 < X < 20) = \int_{15}^{20} 0.2e^{-0.2x} dx = -e^{-0.2x} \Big|_{15}^{20} \approx 0.03154.$$

(c) By the memoryless property, we have

$$P(X > 15+5|X > 5) = P(X > 15) = \int_{15}^{\infty} 0.2e^{-0.2x} dx = -e^{-0.2x} \Big|_{15}^{\infty} \approx 0.04985 \blacksquare$$

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 $g(x) = P(X > x)$. Since X is memoryless then

$$P(X > t) = P(X > s+t|X > s) = \frac{P(X > s+t \text{ and } X > s)}{P(X > s)} = \frac{P(X > s+t)}{P(X > s)}$$

and this implies

$$P(X > s+t) = P(X > s)P(X > t)$$

Hence, g satisfies the equation

$$g(s+t) = g(s)g(t).$$

Theorem 26.1

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 $[g(\frac{1}{n})]^n = g(\frac{1}{n})g(\frac{1}{n}) \cdots g(\frac{1}{n}) = g(\frac{n}{n}) = c$. Thus, $g(\frac{1}{n}) = c^{\frac{1}{n}}$.

Next, let m and n be two positive integers. Then $g(\frac{m}{n}) = g(m \cdot \frac{1}{n}) = g(\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}) = [g(\frac{1}{n})]^m = c^{\frac{m}{n}}$.

Now, if t is a positive real number then we can find a sequence t_n of positive rational numbers such that $\lim_{n \rightarrow \infty} t_n = t$. (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 $F(x) = P(X \leq x) = 1 - e^{-\lambda x}$ and hence $f(x) = F'(x) = \lambda e^{-\lambda x}$ which shows that X is exponentially distributed.

Example 26.6

You call the customer support department of a certain company and are placed on hold. Suppose the amount of time until a service agent assists you has an exponential distribution with mean 5 minutes. Given that you have already been on hold for 2 minutes, what is the probability that you 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 $\lambda = \frac{1}{5}$. Then

$$P(X > 3 + 2 | X > 2) = P(X > 3) = 1 - F(3) = e^{-\frac{3}{5}} \blacksquare$$

Example 26.7

Suppose that the duration of a phone call (in minutes) is an exponential random variable X with $\lambda = 0.1$.

- (a) What is the probability that a given phone call lasts more than 10 minutes?
- (b) Suppose we know that a phone call has already lasted 10 minutes. What is the probability that it last at least 10 more minutes?

Solution.

(a)

$$P(X > 10) = 1 - F(10) = e^{-1} \approx 0.368$$

(b)

$$P(X > 10 + 10 | X > 10) = P(X > 10) \approx 0.368 \blacksquare$$

Problems

Problem 26.1

Let X have an exponential distribution with a mean of 40. Compute $P(X < 36)$.

Problem 26.2

Let X be an exponential function with mean equals to 5. Graph $f(x)$ and $F(x)$.

Problem 26.3

The lifetime (measured in years) of a radioactive element is a continuous random variable with the following p.d.f.:

$$f(x) = \begin{cases} \frac{1}{100}e^{-\frac{x}{100}} & t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

What is the probability that an atom of this element will decay within 50 years?

Problem 26.4

The average number of radioactive particles passing through a counter during 1 millisecond in a lab experiment is 4. What is the probability that more than 2 milliseconds pass between particles?

Problem 26.5

The life length of a computer is exponentially distributed with mean 5 years. You bought an old (working) computer for 10 dollars. What is the probability that it would still work for more than 3 years?

Problem 26.6

Suppose the wait time X for service at the post office has an exponential distribution with mean 3 minutes. A customer first in line will be served immediately.

(a) If you enter the post office immediately behind another customer, what is the probability you wait over 5 minutes?

(b) Under the same conditions, what is the probability of waiting between 2 and 4 minutes?

Problem 26.7

During busy times buses arrive at a bus stop about every three minutes, so if we measure x in minutes the rate λ of the exponential distribution is $\lambda = \frac{1}{3}$.

- (a) What is the probability of having to wait 6 or minutes for the bus?
- (b) What is the probability of waiting between 4 and 7 minutes for a bus?
- (c) What is the probability of having to wait at least 9 more minutes for the bus given that you have already waited 3 minutes?

Problem 26.8

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 26.9

The lifetime (in hours) of a lightbulb is an exponentially distributed random variable with parameter $\lambda = 0.01$.

- (a) What is the probability that the light bulb is still burning one week after it is installed?
- (b) Assume that the bulb was installed at noon today and assume that at 3:00pm tomorrow you notice that the bulb is still working. What is the chance that the bulb will burn out at some time between 4:30pm and 6:00pm tomorrow?

Problem 26.10 †

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 26.11 †

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 26.12 ‡

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 26.13 ‡

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 26.14 ‡

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.

Problem 26.15 ‡

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 26.16

Let X be an exponential random variable such that $P(X \leq 2) = 2P(X > 4)$. Find the variance of X .

Problem 26.17

Customers arrive randomly and independently at a service window, and the time between arrivals has an exponential distribution with a mean of 12 minutes. Let X equal the number of arrivals per hour. What is $P(X = 10)$?
Hint: When the time between successive arrivals has an exponential distribution with mean $\frac{1}{\alpha}$ (units of time), then the number of arrivals per unit time has a Poisson distribution with parameter (mean) α .

27 Gamma and Beta Distributions

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,

$$\Gamma(1) = \int_0^{\infty} e^{-y} dy = -e^{-y} \Big|_0^{\infty} = 1.$$

For $\alpha > 1$ we can use integration by parts with $u = y^{\alpha-1}$ and $dv = e^{-y} dy$ to obtain

$$\begin{aligned} \Gamma(\alpha) &= -e^{-y} y^{\alpha-1} \Big|_0^{\infty} + \int_0^{\infty} e^{-y} (\alpha-1) y^{\alpha-2} dy \\ &= (\alpha-1) \int_0^{\infty} e^{-y} y^{\alpha-2} dy \\ &= (\alpha-1) \Gamma(\alpha-1) \end{aligned}$$

If n is a positive integer greater than 1 then by applying the previous relation repeatedly we find

$$\begin{aligned} \Gamma(n) &= (n-1) \Gamma(n-1) \\ &= (n-1)(n-2) \Gamma(n-2) \\ &= \dots \\ &= (n-1)(n-2) \dots 3 \cdot 2 \cdot \Gamma(1) = (n-1)! \end{aligned}$$

A **Gamma** random variable with parameters $\alpha > 0$ and $\lambda > 0$ has a pdf

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x} (\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

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 $x = \lambda y$.

Note that the above computation involves a $\Gamma(r)$ integral. Thus, the origin of the name of the random variable.

Figure 27.1 shows some examples of Gamma distribution.

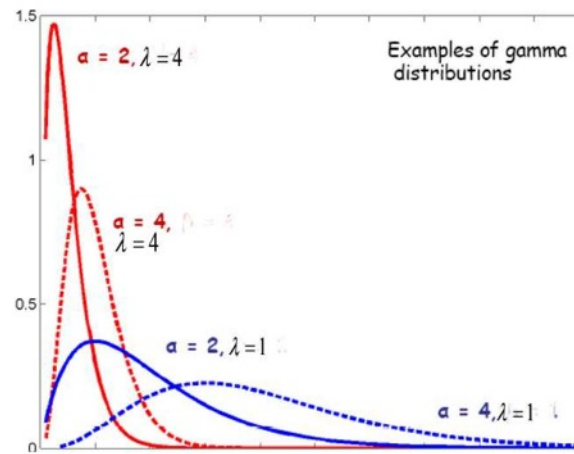


Figure 27.1

Theorem 27.1

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)

$$\begin{aligned}
E(X^2) &= \frac{1}{\Gamma(\alpha)} \int_0^{\infty} x^2 e^{-\lambda x} \lambda^\alpha x^{\alpha-1} dx \\
&= \frac{1}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha+1} \lambda^\alpha e^{-\lambda x} dx \\
&= \frac{\Gamma(\alpha+2)}{\lambda^2 \Gamma(\alpha)} \int_0^{\infty} \frac{x^{\alpha+1} \lambda^{\alpha+2} e^{-\lambda x}}{\Gamma(\alpha+2)} dx \\
&= \frac{\Gamma(\alpha+2)}{\lambda^2 \Gamma(\alpha)}
\end{aligned}$$

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,

$$Var(X) = E(X^2) - (E(X))^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2} \blacksquare$$

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 $(\alpha, \lambda) = (\frac{1}{2}, \frac{n}{2})$ where n is a positive integer. This distribution is called the **chi-squared** distribution. The gamma random variable can be used to model the waiting time until a number of random events occurs. The number of random events sought is α in the formula of $f(x)$.

Example 27.1

In a certain city, the daily consumption of electric power in millions of kilowatt hours 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 power plant of this city has a daily capacity of 12 million kWh, what is the probability that this power supply will be inadequate on a given day? Set up the appropriate integral but do not evaluate.

Solution.

(a) The random variable is the daily consumption of power in kilowatt hours. 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 $\frac{1}{2^3\Gamma(3)} \int_{12}^{\infty} x^2 e^{-\frac{x}{2}} dx = \frac{1}{16} \int_{12}^{\infty} x^2 e^{-\frac{x}{2}} dx$ ■

Problems

Problem 27.1

Let X be a Gamma random variable with $\alpha = 4$ and $\lambda = \frac{1}{2}$. Compute $P(2 < X < 4)$.

Problem 27.2

If X has a probability density function given by

$$f(x) = \begin{cases} 4x^2e^{-2x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find the mean and the variance.

Problem 27.3

Let X be a gamma random variable with $\lambda = 1.8$ and $\alpha = 3$. Compute $P(X > 3)$.

Problem 27.4

A fisherman whose average time for catching a fish is 35 minutes wants to bring home exactly 3 fishes. What is the probability he will need between 1 and 2 hours to catch them?

Problem 27.5

Suppose the time (in hours) taken by a professor to grade a complicated exam 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 grade an exam?

Problem 27.6

Suppose the continuous random variable X has the following pdf:

$$f(x) = \begin{cases} \frac{1}{16}x^2e^{-\frac{x}{2}} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find $E(X^3)$.

Problem 27.7

Let X be the standard normal distribution. Show that X^2 is a gamma distribution with $\alpha = \beta = \frac{1}{2}$.

Problem 27.8

Let X be a gamma random variable with parameter (α, λ) . Find $E(e^{tX})$.

Problem 27.9

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 27.10

If a company employs n salespersons, its gross sales in thousands of dollars may be regarded as a random variable having a gamma distribution with $\alpha = 80\sqrt{n}$ and $\lambda = \frac{1}{2}$. If the sales cost \$8,000 per salesperson, how many salesperson should the company employ to maximize the expected profit?

Problem 27.11

Lifetimes of electrical components follow a Gamma distribution with $\alpha = 3$, and $\lambda = \frac{1}{6}$.

(a) Give the density function (be very specific to any numbers), as well as the mean and standard deviation of the lifetimes.

(b) The monetary value to the firm of a component with a lifetime of X is $V = 3X^2 + X - 1$. Give the expected monetary value of a component.

The Beta Distribution

A random variable is said to have a **beta** distribution with parameters $a > 0$ and $b > 0$ if its density is given by

$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

where

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$$

is the beta function.

Example 27.2

Verify that the integral of the beta density function with parameters (2, 4) from $-\infty$ to ∞ equals 1.

Solution.

Using integration by parts, we have

$$\int_{-\infty}^{\infty} f(x) dx = 20 \int_0^1 x(1-x)^3 dx = 20 \left[-\frac{1}{4}x(1-x)^4 - \frac{1}{20}(1-x)^5 \right]_0^1 = 1 \blacksquare$$

Since the support of X is $0 < x < 1$, the beta distribution is a popular probability model for proportions. The following theorem provides a relationship between the gamma and beta functions.

Theorem 27.2

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Proof.

We have

$$\begin{aligned} \Gamma(a+b)B(a, b) &= \int_0^{\infty} t^{a+b-1} e^{-t} dt \int_0^1 x^{a-1} (1-x)^{b-1} dx \\ &= \int_0^{\infty} t^{b-1} e^{-t} dt \int_0^t u^{a-1} \left(1 - \frac{u}{t}\right)^{b-1} du, \quad u = xt \\ &= \int_0^{\infty} e^{-t} dt \int_0^t u^{a-1} (t-u)^{b-1} du \\ &= \int_0^{\infty} u^{a-1} du \int_u^{\infty} e^{-t} (t-u)^{b-1} dt \\ &= \int_0^{\infty} u^{a-1} e^{-u} du \int_0^{\infty} e^{-v} v^{b-1} dv, \quad v = t-u \\ &= \Gamma(a)\Gamma(b) \blacksquare \end{aligned}$$

Theorem 27.3

If X is a beta random variable with parameters a and b then

- (a) $E(X) = \frac{a}{a+b}$
 (b) $Var(X) = \frac{ab}{(a+b)^2(a+b+1)}$.

Proof.

(a)

$$\begin{aligned} E(X) &= \frac{1}{B(a, b)} \int_0^1 x^a (1-x)^{b-1} dx \\ &= \frac{B(a+1, b)}{B(a, b)} = \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \\ &= \frac{a\Gamma(a)\Gamma(a+b)}{(a+b)\Gamma(a+b)\Gamma(a)} \\ &= \frac{a}{a+b} \end{aligned}$$

(b)

$$\begin{aligned} E(X^2) &= \frac{1}{B(a, b)} \int_0^1 x^{a+1} (1-x)^{b-1} dx = \frac{B(a+2, b)}{B(a, b)} = \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \\ &= \frac{a(a+1)\Gamma(a)\Gamma(a+b)}{(a+b)(a+b+1)\Gamma(a+b)\Gamma(a)} = \frac{a(a+1)}{(a+b)(a+b+1)} \end{aligned}$$

Hence,

$$Var(X) = E(X^2) - (E(X))^2 = \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2} = \frac{ab}{(a+b)^2(a+b+1)} \blacksquare$$

Example 27.3

The proportion of the gasoline supply that is sold during the week is a beta random variable with $a = 4$ and $b = 2$. What is the probability of selling at least 90% of the stock in a given week?

Solution.

Since

$$B(4, 2) = \frac{\Gamma(4)\Gamma(2)}{\Gamma(6)} = \frac{3!1!}{5!} = \frac{1}{20}$$

the density function is $f(x) = 20x^3(1-x)$. Thus,

$$P(X > 0.9) = 20 \int_{0.9}^1 x^3(1-x) dx = 20 \left[\frac{x^4}{4} - \frac{x^5}{5} \right]_{0.9}^1 \approx 0.08146 \blacksquare$$

Example 27.4

Let

$$f(x) = \begin{cases} 12x^2(1-x) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the mean and the variance.

Solution.

Here, $f(x)$ is a beta density function with $a = 3$ and $b = 2$. Thus, $E(X) = \frac{a}{a+b} = \frac{3}{5}$ and $\text{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)} = \frac{1}{25}$ ■

Problems

Problem 27.12

Let

$$f(x) = \begin{cases} kx^3(1-x)^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Find the value of k that makes $f(x)$ a density function.
- Find the mean and the variance of X .

Problem 27.13

Suppose that X has a probability density function given by

$$f(x) = \begin{cases} 6x(1-x) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Find $F(x)$.
- Find $P(0.5 < X < 0.8)$.

Problem 27.14

A management firm handles investment accounts for a large number of clients. The percent of clients who telephone the firm for information or services in a given month is a beta random variable with $a = 4$ and $b = 3$.

- Find the density function $f(x)$ of X .
- Find the cumulative density function $F(x)$.

Problem 27.15

A company markets a new product and surveys customers on their satisfaction with their product. The fraction of customers who are dissatisfied has a beta distribution with $a = 2$ and $b = 4$. What is the probability that no more than 30% of the customers are dissatisfied?

Problem 27.16

Consider the following hypothetical situation. Grade data indicates that on the average 27% of the students in senior engineering classes have received A grades. There is variation among classes, however, and the proportion X must be considered a random variable. From past data we have measured a standard deviation of 15%. We would like to model the proportion X of A grades with a Beta distribution.

- Find the density function of X .
- Find the probability that more than 50% of the students had an A.

Problem 27.17

Let X be a beta random variable with density function $f(x) = kx^2(1 - x)$, $0 < x < 1$ and 0 otherwise.

- (a) Find the value of k .
 (b) Find $P(0.5 < X < 1)$.

Problem 27.18

Let X be a beta distribution with parameters a and b . Show that

$$E(X^n) = \frac{B(a+n, b)}{B(a, b)}.$$

Problem 27.19

Suppose that X is a beta distribution with parameters (a, b) . Show that $Y = 1 - X$ is a beta distribution with parameters (b, a) . Hint: Find $F_Y(y)$ and then differentiate with respect to y .

Problem 27.20

If the annual proportion of new restaurants that fail in a given city may be looked upon as a random variable having a beta distribution with $a = 1$ and $b = 5$, find

- (a) the mean of this distribution, that is, the annual proportion of new restaurants that can be expected to fail in a given city;
 (b) the probability that at least 25 percent of all new restaurants will fail in the given city in any one year.

Problem 27.21

If the annual proportion of erroneous income tax returns filed with the IRS can be looked upon as a random variable having a beta distribution with parameters $(2, 9)$, what is the probability that in any given year there will be fewer than 10 percent erroneous returns?

Problem 27.22

Suppose that the continuous random variable X has the density function

$$f(x) = \begin{cases} \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

where

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx, a > 0, b > 0.$$

If $a = 5$ and $b = 6$, what is $E[(1 - X)^{-4}]$?

Problem 27.23

The amount of impurity in batches of a chemical follow a Beta distribution with $a = 1$, and $b = 2$.

- (a) Give the density function, as well as the mean amount of impurity.
- (b) If the amount of impurity exceeds 0.7, the batch cannot be sold. What proportion of batches can be sold?

28 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 28.1

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

$$\begin{aligned} F(y) &= P(Y \leq y) \\ &= P(X^3 \leq y) \\ &= P(X \leq y^{\frac{1}{3}}) \\ &= \int_0^{y^{\frac{1}{3}}} 6x(1-x) dx \\ &= 3y^{\frac{2}{3}} - 2y \end{aligned}$$

Hence, $f(y) = F'(y) = 2(y^{-\frac{1}{3}} - 1)$, for $0 < y < 1$ and 0 elsewhere ■

Example 28.2

Let X be a random variable with probability density $f(x)$. Find the probability density function of $Y = |X|$.

Solution.

Clearly, $F(y) = 0$ for $y \leq 0$. So assume that $y > 0$. Then

$$\begin{aligned} F(y) &= P(Y \leq y) \\ &= P(|X| \leq y) \\ &= P(-y \leq X \leq y) \\ &= F(y) - F(-y) \end{aligned}$$

Thus, $f(y) = F'(y) = f(y) + f(-y)$ for $y > 0$ and 0 elsewhere ■

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 28.1

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

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\ &= P(X \leq g^{-1}(y)) = F_X(g^{-1}(y)) \end{aligned}$$

Differentiating we find

$$f_Y(y) = \frac{dF_Y(y)}{dy} = f_X[g^{-1}(y)] \frac{d}{dy} g^{-1}(y).$$

Now, suppose that $g(\cdot)$ is decreasing. Then

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\ &= P(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y)) \end{aligned}$$

Differentiating we find

$$f_Y(y) = \frac{dF_Y(y)}{dy} = -f_X[g^{-1}(y)] \frac{d}{dy} g^{-1}(y) \blacksquare$$

Example 28.3

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 28.4

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 28.5

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 $x \leq 0$. Now, for $x > 0$ we have

$$F_{|X|}(x) = P(|X| \leq x) = \int_{-x}^x \frac{1}{\pi(x^2 + 1)} dx = \frac{2}{\pi} \tan^{-1} x.$$

Hence,

$$F_{|X|}(x) = \begin{cases} 0 & x \leq 0 \\ \frac{2}{\pi} \tan^{-1} x & x > 0 \end{cases}$$

(b) X^2 also takes only positive values, so the density $f_{X^2}(x) = 0$ for $x \leq 0$. Furthermore, $F_{X^2}(x) = P(X^2 \leq x) = P(|X| \leq \sqrt{x}) = \frac{2}{\pi} \tan^{-1} \sqrt{x}$. So by differentiating we get

$$f_{X^2}(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{\pi\sqrt{x}(1+x)} & x > 0 \end{cases} \blacksquare$$

Remark 28.1

In general, if a function does not have a unique inverse, we must sum over all possible inverse values.

Example 28.6

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,

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Differentiate both sides to obtain,

$$f_Y(y) = \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}} \blacksquare$$

Problems

Problem 28.1

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 28.2

A process for refining sugar yields up to 1 ton of pure sugar per day, but the actual amount produced, is a random variable because of machine breakdowns and other slowdowns. Suppose that X has the density function given by

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

The company is paid at the rate of \$300 per ton for the refined sugar, but it also has a fixed overhead cost of \$100 per day. Thus, the daily profit, in hundreds of dollars, is $Y = 3X - 1$. Find probability density function for Y .

Problem 28.3

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 28.4

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 distribution of $Y = e^X$?

Problem 28.5

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 28.6

Let X be a random variable that is uniformly distributed in $(0,1)$. Find the probability density function of $Y = -\ln X$.

Problem 28.7

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 28.8

Suppose X has the uniform distribution on $(0, 1)$. Compute the probability density function and expected value of:

- (a) X^α , $\alpha > 0$ (b) $\ln X$ (c) e^X (d) $\sin \pi X$

Problem 28.9 †

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?

Problem 28.10 †

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 28.11 †

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 28.12 ‡

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 28.13 ‡

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 28.14 ‡

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 28.15

Let X have normal distribution with mean 1 and standard deviation 2.

(a) Find $P(|X| \leq 1)$.

(b) Let $Y = e^X$. Find the probability density function $f_Y(y)$ of Y .

Problem 28.16

Let X be a uniformly distributed random variable on the interval $(-1, 1)$. Show that $Y = X^2$ is a beta random variable with parameters $(\frac{1}{2}, 1)$.

Problem 28.17

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 28.18

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 28.19

If $f(x) = xe^{-\frac{x^2}{2}}$, for $x > 0$ and $Y = \ln X$, find the density function for Y .

Problem 28.20

A supplier faces a daily demand (Y , in proportion of her daily supply) with the following density function.

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

- (a) Her profit is given by $Y = 10X - 2$. Find the density function of her daily profit.
- (b) What is her average daily profit?
- (c) What is the probability she will lose money on a day?

Joint Distributions

There are many situations which involve the presence of several random variables and we are interested in their joint behavior. For example:

- (i) A meteorological station may record the wind speed and direction, air pressure and the air temperature.
- (ii) Your physician may record your height, weight, blood pressure, cholesterol level and more.

This chapter is concerned with the joint probability structure of two or more random variables defined on the same sample space.

29 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

$$F_{XY}(x, y) = P(X \leq x, Y \leq y) = P(\{e \in S : X(e) \leq x \text{ and } Y(e) \leq y\}).$$

Example 29.1

Consider the experiment of throwing a fair coin and a fair die simultaneously. The sample space is

$$S = \{(H, 1), (H, 2), \dots, (H, 6), (T, 1), (T, 2), \dots, (T, 6)\}.$$

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.

$$\begin{aligned}
 F_{XY}(1, 2) &= P(X \leq 1, Y \leq 2) \\
 &= P(\{(H, 1), (H, 2), (T, 1), (T, 2)\}) \\
 &= \frac{4}{12} = \frac{1}{3} \blacksquare
 \end{aligned}$$

In what follows, individual cdfs will be referred to as **marginal distributions**. These cdfs are obtained from the joint cumulative distribution as follows

$$\begin{aligned}
 F_X(x) &= P(X \leq x) \\
 &= P(X \leq x, Y < \infty) \\
 &= P(\lim_{y \rightarrow \infty} \{X \leq x, Y \leq y\}) \\
 &= \lim_{y \rightarrow \infty} P(X \leq x, Y \leq y) \\
 &= \lim_{y \rightarrow \infty} F_{XY}(x, y) = F_{XY}(x, \infty)
 \end{aligned}$$

In a similar way, one can show that

$$F_Y(y) = \lim_{x \rightarrow \infty} F_{XY}(x, y) = F_{XY}(\infty, y).$$

It is easy to see that

$$F_{XY}(\infty, \infty) = P(X < \infty, Y < \infty) = 1.$$

Also,

$$F_{XY}(-\infty, y) = 0.$$

This follows from

$$0 \leq F_{XY}(-\infty, y) = P(X < -\infty, Y \leq y) \leq P(X < -\infty) = F_X(-\infty) = 0.$$

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,

$$\begin{aligned}
 P(X > x, Y > y) &= 1 - P(\{X > x, Y > y\}^c) \\
 &= 1 - P(\{X > x\}^c \cup \{Y > y\}^c) \\
 &= 1 - [P(\{X \leq x\} \cup \{Y \leq y\})] \\
 &= 1 - [P(X \leq x) + P(Y \leq y) - P(X \leq x, Y \leq y)] \\
 &= 1 - F_X(x) - F_Y(y) + F_{XY}(x, y)
 \end{aligned}$$

Also, if $a_1 < a_2$ and $b_1 < b_2$ then

$$\begin{aligned}
 P(a_1 < X \leq a_2, b_1 < Y \leq b_2) &= P(X \leq a_2, Y \leq b_2) - P(X \leq a_2, Y \leq b_1) \\
 &\quad - P(X \leq a_1, Y \leq b_2) + P(X \leq a_1, Y \leq b_1) \\
 &= F_{XY}(a_2, b_2) - F_{XY}(a_1, b_2) - F_{XY}(a_2, b_1) + F_{XY}(a_1, b_1)
 \end{aligned}$$

This is clear if you use the concept of area shown in Figure 29.1

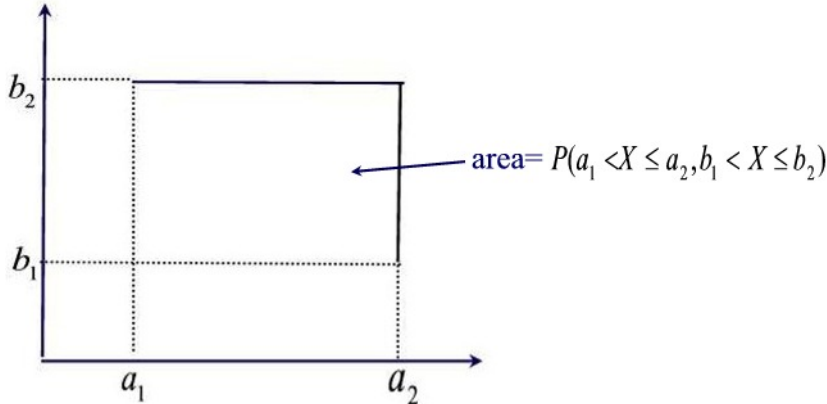


Figure 29.1

If X and Y are both discrete random variables, we define the **joint probability mass function** of X and Y by

$$p_{XY}(x, y) = P(X = x, Y = y).$$

The marginal probability mass function of X can be obtained from $p_{XY}(x, y)$ by

$$p_X(x) = P(X = x) = \sum_{y: p_{XY}(x, y) > 0} p_{XY}(x, y).$$

Similarly, we can obtain the marginal pmf of Y by

$$p_Y(y) = P(Y = y) = \sum_{x:p_{XY}(x,y)>0} p_{XY}(x,y).$$

This simply means 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 29.2

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.

- What is the joint pmf of X and Y ?
- 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?
- What is the joint cdf of X and Y ?
- What is the probability less than 3 heads occur in both the first and last 3 tosses?
- Find the probability that one head appears in the first three tosses.

Solution.

- The joint pmf is given by the following table

$X \setminus Y$	0	1	2	3	$p_X(\cdot)$
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
$p_Y(\cdot)$	2/16	6/16	6/16	2/16	1

- $P((X, Y) \in \{(2, 1), (2, 2), (3, 1), (3, 2)\}) = p(2, 1) + p(2, 2) + p(3, 1) + p(3, 2) = \frac{3}{8}$

- The joint cdf is given by the following table

$X \setminus Y$	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

$$(d) P(X < 3, Y < 3) = F(2, 2) = \frac{13}{16}$$

$$(e) P(X = 1) = P((X, Y) \in \{(1, 0), (1, 1), (1, 2), (1, 3)\}) = 1/16 + 3/16 + 2/16 = \frac{3}{8} \blacksquare$$

Example 29.3

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.

$$p_{XY}(0, 0) = C(2, 2)/C(10, 2) = 1/45$$

$$p_{XY}(0, 1) = C(2, 1)C(5, 1)/C(10, 2) = 10/45$$

$$p_{XY}(0, 2) = C(5, 2)/C(8, 2) = 10/45$$

$$p_{XY}(1, 0) = C(2, 1)C(3, 1)/C(10, 2) = 6/45$$

$$p_{XY}(1, 1) = C(5, 1)C(3, 1)/C(10, 2) = 15/45$$

$$p_{XY}(1, 2) = 0$$

$$p_{XY}(2, 0) = C(3, 2)/C(10, 2) = 3/45$$

$$p_{XY}(2, 1) = 0$$

$$p_{XY}(2, 2) = 0$$

The pmf of X is

$$p_X(0) = P(X = 0) = \sum_{y:p_{XY}(0,y)>0} p_{XY}(0, y) = \frac{1 + 10 + 10}{45} = \frac{21}{45}$$

$$p_X(1) = P(X = 1) = \sum_{y:p_{XY}(1,y)>0} p_{XY}(1, y) = \frac{6 + 15}{45} = \frac{21}{45}$$

$$p_X(2) = P(X = 2) = \sum_{y:p_{XY}(2,y)>0} p_{XY}(2, y) = \frac{3}{45} = \frac{3}{45}$$

The pmf of y is

$$\begin{aligned}
 p_Y(0) = P(Y = 0) &= \sum_{x:p_{XY}(x,0)>0} p_{XY}(x, 0) = \frac{1 + 6 + 3}{45} = \frac{10}{45} \\
 p_Y(1) = P(Y = 1) &= \sum_{x:p_{XY}(x,1)>0} p_{XY}(x, 1) = \frac{10 + 15}{45} = \frac{25}{45} \\
 p_Y(2) = P(Y = 2) &= \sum_{x:p_{XY}(x,2)>0} p_{XY}(x, 2) = \frac{10}{45} = \frac{10}{45} \blacksquare
 \end{aligned}$$

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

$$P((X, Y) \in C) = \iint_{(x,y) \in C} f_{XY}(x, y) dx dy$$

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

$$P(X \in A, Y \in B) = \int_B \int_A f_{XY}(x, y) dx dy$$

As a result of this last equation we can write

$$\begin{aligned}
 F_{XY}(x, y) &= P(X \in (-\infty, x], Y \in (-\infty, y]) \\
 &= \int_{-\infty}^y \int_{-\infty}^x f_{XY}(u, v) du dv
 \end{aligned}$$

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 29.4

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:

$$\begin{aligned} P(X \in A) &= P(X \in A, Y \in (-\infty, \infty)) \\ &= \int_A \int_{-\infty}^{\infty} f_{XY}(x, y) dy dx \\ &= \int_A f_X(x, y) dx \end{aligned}$$

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 29.5

Let X and Y be random variables with joint pdf

$$f_{XY}(x, y) = \begin{cases} \frac{1}{4} & -1 \leq x, y \leq 1 \\ 0 & \text{Otherwise} \end{cases}$$

Determine

- (a) $P(X^2 + Y^2 < 1)$,
- (b) $P(2X - Y > 0)$,
- (c) $P(|X + Y| < 2)$.

Solution.

(a)

$$P(X^2 + Y^2 < 1) = \int_0^{2\pi} \int_0^1 \frac{1}{4} r dr d\theta = \frac{\pi}{4}.$$

(b)

$$P(2X - Y > 0) = \int_{-1}^1 \int_{\frac{y}{2}}^1 \frac{1}{4} dx dy = \frac{1}{2}.$$

Note that $P(2X - Y > 0)$ 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 $-2 < x + y < 2$ then

$$P(|X + Y| < 2) = 1 \blacksquare$$

Joint pdfs and joint cdfs for three or more random variables are obtained as straightforward generalizations of the above definitions and conditions.

Problems

Problem 29.1

A supermarket 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

$X \setminus Y$	0	1	2	3	$p_X(\cdot)$
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
$p_Y(\cdot)$	0.3	0.5	0.125	0.075	1

- Verify that $p_{XY}(x, y)$ is a joint probability mass function.
- Find the probability that more than two customers are in line.
- Find $P(|X - Y| = 1)$, the probability that X and Y differ by exactly 1.
- What is the marginal probability mass function of X ?

Problem 29.2

Two tire-quality experts examine stacks of tires and assign quality ratings to each tire on a 3-point scale. Let X denote the grade given by expert A and let Y denote the grade given by B . The following table gives the joint distribution for X, Y .

$X \setminus Y$	1	2	3	$p_X(\cdot)$
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
$p_Y(\cdot)$	0.23	0.50	0.27	1

Find $P(X \geq 2, Y \geq 3)$.

Problem 29.3

In a randomly chosen lot of 1000 bolts, let X be the number that fail to meet a length specification, and let Y be the number that fail to meet a diameter specification. Assume that the joint probability mass function of X and Y is given in the following table.

$X \setminus Y$	0	1	2	$p_X(\cdot)$
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
$p_Y(\cdot)$	0.65	0.23	0.12	1

- (a) Find $P(X = 0, Y = 2)$.
 (b) Find $P(X > 0, Y \leq 1)$.
 (c) $P(X \leq 1)$.
 (d) $P(Y > 0)$.
 (e) Find the probability that all bolts in the lot meet the length specification.
 (f) Find the probability that all bolts in the lot meet the diameter specification.
 (g) Find the probability that all bolts in the lot meet both specification.

Problem 29.4

Rectangular plastic covers for a compact disc (CD) tray have specifications regarding length and width. Let X and Y be the length and width respectively measured in millimeters. There are six possible ordered pairs (X, Y) and following is the corresponding probability distribution.

$X \setminus Y$	15	16	$p_X(\cdot)$
129	0.12	0.08	0.2
130	0.4	0.30	0.7
131	0.06	0.04	0.1
$p_Y(\cdot)$	0.58	0.42	1

- (a) Find $P(X = 130, Y = 15)$.
 (b) Find $P(X \geq 130, Y \geq 15)$.

Problem 29.5

Suppose the random variables X and Y have a joint pdf

$$f_{XY}(x, y) = \begin{cases} 0.0032(20 - x - y) & 0 \leq x, y \leq 5 \\ 0 & \text{otherwise} \end{cases}$$

Find $P(1 \leq X \leq 2, 2 \leq Y \leq 3)$.

Problem 29.6

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 29.7

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 29.8 ‡

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 29.9 ‡

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 29.10 ‡

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

$$\begin{aligned} P(X = 0, Y = 0) &= \frac{1}{6} \\ P(X = 1, Y = 0) &= \frac{1}{12} \\ P(X = 1, Y = 1) &= \frac{1}{6} \\ P(X = 2, Y = 0) &= \frac{1}{12} \\ P(X = 2, Y = 1) &= \frac{1}{3} \\ P(X = 2, Y = 2) &= \frac{1}{6} \end{aligned}$$

What is the variance of X ?

Problem 29.11 ‡

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 29.12 ‡

Let X and Y be continuous random variables with joint density function

$$f_{XY}(x, y) = \begin{cases} 15y & x^2 \leq y \leq x \\ 0 & \text{otherwise} \end{cases}$$

Find the marginal density function of Y .

Problem 29.13 ‡

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.

Problem 29.14 ‡

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

$$f_{XY}(x, y) = \begin{cases} \frac{1}{64}(10 - xy^2) & 2 \leq x \leq 10, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Calculate the expected age of an insured automobile involved in an accident.

Problem 29.15 ‡

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 29.16 ‡

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 29.17

Suppose the random variables X and Y have a joint pdf

$$f_{XY}(x, y) = \begin{cases} x + y & 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $P(X > \sqrt{Y})$.

Problem 29.18 ‡

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 29.19

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 $0 \leq x \leq 5$ and $0 \leq y \leq 5$. Compute $P(X > 2)$.

Problem 29.20

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.

Problem 29.21

Let X and Y be continuous random variables with joint density function

$$f_{XY}(x, y) = \begin{cases} xy & 0 \leq x \leq 2, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $P(\frac{X}{2} \leq Y \leq X)$.

Problem 29.22

Let X and Y be random variables with common range $\{1, 2\}$ and such that $P(X = 1) = 0.7, P(X = 2) = 0.3, P(Y = 1) = 0.4, P(Y = 2) = 0.6$, and $P(X = 1, Y = 1) = 0.2$.

- (a) Find the joint probability mass function $p_{XY}(x, y)$.
- (b) Find the joint cumulative distribution function $F_{XY}(x, y)$.

30 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

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \quad (30.1)$$

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 30.1

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 30.1 we obtain

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

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

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{y \in B} \sum_{x \in A} p_{XY}(x, y) \\ &= \sum_{y \in B} \sum_{x \in A} p_X(x)p_Y(y) \\ &= \sum_{y \in B} p_Y(y) \sum_{x \in A} p_X(x) \\ &= P(Y \in B)P(X \in A) \end{aligned}$$

and thus equation 30.1 is satisfied. That is, X and Y are independent ■

Example 30.1

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

$Y \setminus X$	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 $P(X \leq 6) = \frac{6}{12} = \frac{1}{2}$, $P(Y = 30) = \frac{4}{12} = \frac{1}{3}$, $P(X \leq 6, Y = 30) = \frac{2}{12} = \frac{1}{6}$. Since, $P(X \leq 6, Y = 30) = P(X \leq 6)P(Y = 30)$, 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 ■

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 30.2

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

$$f_X(x) = \int_0^{\infty} 4e^{-2(x+y)} dy = 2e^{-2x} \int_0^{\infty} 2e^{-2y} dy = 2e^{-2x}, \quad x > 0$$

Similarly, the marginal density $f_Y(y)$ is given by

$$f_Y(y) = \int_0^{\infty} 4e^{-2(x+y)} dx = 2e^{-2y} \int_0^{\infty} 2e^{-2x} dx = 2e^{-2y}, \quad y > 0$$

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 30.3

The joint pdf of X and Y is given by

$$f_{XY}(x, y) = \begin{cases} 3(x+y) & 0 \leq x+y \leq 1, \quad 0 \leq x, y < \infty \\ 0 & \text{Otherwise} \end{cases}$$

Are X and Y independent?

Solution.

For the limit of integration see Figure 30.1 below.

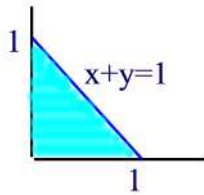


Figure 30.1

The marginal pdf of X is

$$f_X(x) = \int_0^{1-x} 3(x+y) dy = 3xy + \frac{3}{2}y^2 \Big|_0^{1-x} = \frac{3}{2}(1-x^2), \quad 0 \leq x \leq 1$$

The marginal pdf of Y is

$$f_Y(y) = \int_0^{1-y} 3(x+y)dx = \frac{3}{2}x^2 + 3xy \Big|_0^{1-y} = \frac{3}{2}(1-y^2), \quad 0 \leq y \leq 1$$

But

$$f_{XY}(x, y) = 3(x+y) \neq \frac{3}{2}(1-x^2)\frac{3}{2}(1-y^2) = f_X(x)f_Y(y)$$

so that X and Y are dependent ■

Example 30.4

A man and woman decide to meet at a certain location. If each person independently arrives at a time uniformly distributed between 2 and 3 PM, find the probability that the man arrives 10 minutes before the woman.

Solution.

Let X represent the number of minutes past 2 the man arrives and Y the number of minutes past 2 the woman arrives. X and Y are independent random variables each uniformly distributed over $(0,60)$. Then

$$\begin{aligned} P(X + 10 < Y) &= \iint_{x+10 < y} f_{XY}(x, y) dx dy = \iint_{x+10 < y} f_X(x) f_Y(y) dx dy \\ &= \int_{10}^{60} \int_0^{y-10} \left(\frac{1}{60}\right)^2 dx dy \\ &= \left(\frac{1}{60}\right)^2 \int_{10}^{60} (y-10) dy \approx 0.347 \quad \blacksquare \end{aligned}$$

The following theorem provides a necessary and sufficient condition for two random variables to be independent.

Theorem 30.2

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

$$\begin{aligned} CD &= \left(\int_{-\infty}^{\infty} h(x)dx \right) \left(\int_{-\infty}^{\infty} g(y)dy \right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x)g(y)dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y)dx dy = 1 \end{aligned}$$

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 30.5

The joint pdf of X and Y is given by

$$f_{XY}(x, y) = \begin{cases} xye^{-\frac{(x^2+y^2)}{2}} & 0 \leq x, y < \infty \\ 0 & \text{Otherwise} \end{cases}$$

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 30.6

The joint pdf of X and Y is given by

$$f_{XY}(x, y) = \begin{cases} x + y & 0 \leq x, y < 1 \\ 0 & \text{Otherwise} \end{cases}$$

Are X and Y independent?

Solution.

Let

$$I(x, y) = \begin{cases} 1 & 0 \leq x < 1, 0 \leq y < 1 \\ 0 & \text{otherwise} \end{cases}$$

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 30.7

Let X and Y be two independent random variable with X having a normal distribution with mean μ and variance 1 and Y being the standard normal distribution.

- Find the density of $Z = \min\{X, Y\}$.
- For each $t \in \mathbb{R}$ calculate $P(\max(X, Y) - \min(X, Y) > t)$.

Solution.

- Fix a real number z . Then

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = 1 - P(\min(X, Y) > z) \\ &= 1 - P(X > z)P(Y > z) = 1 - (1 - \Phi(z - \mu))(1 - \Phi(z)) \end{aligned}$$

Hence,

$$f_Z(z) = (1 - \Phi(z - \mu))\phi(z) + (1 - \Phi(z))\phi(z - \mu).$$

- If $t \leq 0$ then $P(\max(X, Y) - \min(X, Y) > t) = 1$. If $t > 0$ then

$$\begin{aligned} P(\max(X, Y) - \min(X, Y) > t) &= P(|X - Y| > t) \\ &= 1 - \Phi\left(\frac{t - \mu}{\sqrt{2}}\right) + \Phi\left(\frac{-t - \mu}{\sqrt{2}}\right) \end{aligned}$$

Note that $X - Y$ is normal with mean μ and variance 2 ■

Example 30.8

Suppose X_1, \dots, X_n are independent and identically distributed random variables with cdf $F_X(x)$. Define U and L as

$$U = \max\{X_1, X_2, \dots, X_n\}$$

$$L = \min\{X_1, X_2, \dots, X_n\}$$

- (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

$$\{U \leq u\} \Leftrightarrow \{X_1 \leq u, X_2 \leq u, \dots, X_n \leq u\}.$$

Thus,

$$\begin{aligned} F_U(u) &= P(U \leq u) = P(X_1 \leq u, X_2 \leq u, \dots, X_n \leq u) \\ &= P(X_1 \leq u)P(X_2 \leq u) \cdots P(X_n \leq u) = (F(X))^n \end{aligned}$$

- (b) Note the following equivalence of events

$$\{L > l\} \Leftrightarrow \{X_1 > l, X_2 > l, \dots, X_n > l\}.$$

Thus,

$$\begin{aligned} F_L(l) &= P(L \leq l) = 1 - P(L > l) \\ &= 1 - P(X_1 > l, X_2 > l, \dots, X_n > l) \\ &= 1 - P(X_1 > l)P(X_2 > l) \cdots P(X_n > l) \\ &= 1 - (1 - F(X))^n \end{aligned}$$

- (c) No. First note that $P(L > l) = 1 - F_L(l)$. From the definition of cdf there must be a number l_0 such that $F_L(l_0) \neq 1$. Thus, $P(L > l_0) \neq 0$. But $P(L > l_0 | U \leq u) = 0$ for any $u < l_0$. This shows that $P(L > l_0 | U \leq u) \neq P(L > l_0)$ ■

Problems

Problem 30.1

Let X and Y be random variables with joint pdf given by

$$f_{XY}(x, y) = \begin{cases} e^{-(x+y)} & 0 \leq x, y \\ 0 & \text{otherwise} \end{cases}$$

- (a) Are X and Y independent?
- (b) Find $P(X < Y)$.
- (c) Find $P(X < a)$.

Problem 30.2

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 $R = \{(x, y) : -1 \leq x \leq 1, -1 \leq y \leq 1\}$. Show that X and Y are independent, with each being distributed uniformly over $(-1, 1)$.
- (c) Find $P(X^2 + Y^2 \leq 1)$.

Problem 30.3

Let X and Y be random variables with joint pdf given by

$$f_{XY}(x, y) = \begin{cases} 6(1-y) & 0 \leq x \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Find $P(X \leq \frac{3}{4}, Y \geq \frac{1}{2})$.
- (b) Find $f_X(x)$ and $f_Y(y)$.
- (c) Are X and Y independent?

Problem 30.4

Let X and Y have the joint pdf given by

$$f_{XY}(x, y) = \begin{cases} kxy & 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Find k .
- (b) Find $f_X(x)$ and $f_Y(y)$.
- (c) Are X and Y independent?

Problem 30.5

Let X and Y have joint density

$$f_{XY}(x, y) = \begin{cases} kxy^2 & 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Find k .
- (b) Compute the marginal densities of X and of Y .
- (c) Compute $P(Y > 2X)$.
- (d) Compute $P(|X - Y| < 0.5)$.
- (e) Are X and Y independent?

Problem 30.6

Suppose the joint density of random variables X and Y is given by

$$f_{XY}(x, y) = \begin{cases} kx^2y^{-3} & 1 \leq x, y \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Find k .
- (b) Are X and Y independent?
- (c) Find $P(X > Y)$.

Problem 30.7

Let X and Y be continuous random variables, with the joint probability density function

$$f_{XY}(x, y) = \begin{cases} \frac{3x^2+2y}{24} & 0 \leq x, y \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Find $f_X(x)$ and $f_Y(y)$.
- (b) Are X and Y independent?
- (c) Find $P(X + 2Y < 3)$.

Problem 30.8

Let X and Y have joint density

$$f_{XY}(x, y) = \begin{cases} \frac{4}{9} & x \leq y \leq 3 - x, 0 \leq x \\ 0 & \text{otherwise} \end{cases}$$

- (a) Compute the marginal densities of X and Y .
- (b) Compute $P(Y > 2X)$.
- (c) Are X and Y independent?

Problem 30.9 ‡

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 30.10 ‡

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 30.11 ‡

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 30.12 ‡

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 30.13 ‡

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 30.14 ‡

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 30.15 ‡

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 30.16 ‡

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 30.17

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 $P(X^2 \geq Y^3)$?

Problem 30.18

Suppose that discrete random variables X and Y each take only the values 0 and 1. It is known that $P(X = 0|Y = 1) = 0.6$ and $P(X = 1|Y = 0) = 0.7$. Is it possible that X and Y are independent? Justify your conclusion.

31 Sum of Two Independent Random Variables

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.

31.1 Discrete Case

In this subsection we consider only sums of discrete random variables, reserving the case of continuous random variables for the next subsection. 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

$$P(X + Y = n) = \sum_{k=0}^n P(X = k)P(Y = n - k).$$

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 31.1

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,

$$P(Z = 2) = p_X(1)p_X(1) = \frac{1}{36}$$

$$P(Z = 3) = p_X(1)p_X(2) + p_X(2)p_X(1) = \frac{2}{36}$$

$$P(Z = 4) = p_X(1)p_X(3) + p_X(2)p_X(2) + p_X(3)p_X(1) = \frac{3}{36}$$

Continuing in this way we would find $P(Z = 5) = 4/36$, $P(Z = 6) = 5/36$, $P(Z = 7) = 6/36$, $P(Z = 8) = 5/36$, $P(Z = 9) = 4/36$, $P(Z = 10) = 3/36$, $P(Z = 11) = 2/36$, and $P(Z = 12) = 1/36$ ■

Example 31.2

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 $0 \leq k \leq n$. Moreover, $A_i \cap A_j = \emptyset$ for $i \neq j$. Thus,

$$\begin{aligned} p_{X+Y}(n) &= P(X + Y = n) = \sum_{k=0}^n P(X = k, Y = n - k) \\ &= \sum_{k=0}^n P(X = k)P(Y = n - k) \\ &= \sum_{k=0}^n e^{-\lambda_1} \frac{\lambda_1^k}{k!} e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} \\ &= e^{-(\lambda_1 + \lambda_2)} \sum_{k=0}^n \frac{\lambda_1^k \lambda_2^{n-k}}{k!(n-k)!} \\ &= \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} \lambda_1^k \lambda_2^{n-k} \\ &= \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} (\lambda_1 + \lambda_2)^n \end{aligned}$$

Thus, $X + Y$ is a Poisson random variable with parameters $\lambda_1 + \lambda_2$ ■

Example 31.3

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 31.4

Alice and Bob flip bias coins independently. Alice's coin comes up heads with probability $1/4$, while Bob's coin comes up head with probability $3/4$. Each stop as soon as they get a head; that is, Alice stops when she gets a head while Bob stops when he gets a head. What is the pmf of the total amount of flips until both stop? (That is, what is the pmf of the combined total amount of flips for both Alice and Bob until they stop?)

Solution.

Let X and Y be the number of flips until Alice and Bob stop, respectively. Thus, $X + Y$ is the total number of flips until both stop. The random variables X and Y are independent geometric random variables with parameters $1/4$ and $3/4$, respectively. By convolution, we have

$$\begin{aligned} p_{X+Y}(n) &= \sum_{k=1}^{n-1} \frac{1}{4} \left(\frac{3}{4}\right)^{k-1} \frac{3}{4} \left(\frac{1}{4}\right)^{n-k-1} \\ &= \frac{1}{4^n} \sum_{k=1}^{n-1} 3^k = \frac{3}{2} \frac{3^{n-1} - 1}{4^n} \quad \blacksquare \end{aligned}$$

Problems

Problem 31.1

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 31.2

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 31.3

Let X and Y be independent random variables each geometrically distributed with parameter p , i.e.

$$p_X(n) = p_Y(n) = \begin{cases} p(1-p)^{n-1} & n = 1, 2, \dots \\ 0 & \text{otherwise} \end{cases}$$

Find the probability mass function of $X + Y$.

Problem 31.4

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 31.5 †

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 31.6

Suppose X and Y are independent, each having Poisson distribution with means 2 and 3, respectively. Let $Z = X + Y$. Find $P(X + Y = 1)$.

Problem 31.7

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) $P(X + Y = 2)$
 (b) $P(Y > X)$

Problem 31.8

An insurance company has two clients. The random variables representing the claims filed by each client are X and Y . X and Y are independent with common pmf

x	0	1	2	y
$p_X(x)$	0.5	0.25	0.25	$p_Y(y)$

Find the probability density function of $X + Y$.

Problem 31.9

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 density function of $X + Y$.

Problem 31.10

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 31.11

Let X, Y, Z be independent Poisson random variables with $E(X) = 3$, $E(Y) = 1$, and $E(Z) = 4$. What is $P(X + Y + Z \leq 1)$?

Problem 31.12

If the number of typographical errors per page type by a certain typist follows a Poisson distribution with a mean of λ , find the probability that the total number of errors in 10 randomly selected pages is 10.

31.2 Continuous Case

In this subsection we consider the continuous version of the problem posed in Section 31.1: How are sums of independent continuous random variables distributed?

Example 31.5

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 31.1, we get

$$F_Z(a) = P(Z \leq a) = \int_0^a \int_0^{a-y} 6e^{-3x-2y} dx dy = 1 + 2e^{-3a} - 3e^{-2a}$$

and differentiating with respect to a we find

$$f_Z(a) = 6(e^{-2a} - e^{-3a})$$

for $a > 0$ and 0 elsewhere ■

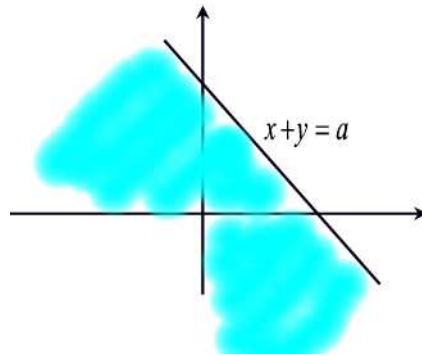


Figure 31.1

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 31.1

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:

$$\begin{aligned}F_{X+Y}(a) &= P(X + Y \leq a) = \iint_{x+y \leq a} f_X(x)f_Y(y)dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x)f_Y(y)dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x)dx f_Y(y)dy \\ &= \int_{-\infty}^{\infty} F_X(a - y)f_Y(y)dy\end{aligned}$$

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 31.6

Let X and Y be two independent random variables uniformly distributed on $[0, 1]$. Compute the distribution of $X + Y$.

Solution.

Since

$$f_X(a) = f_Y(a) = \begin{cases} 1 & 0 \leq a \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

by the previous theorem

$$f_{X+Y}(a) = \int_0^1 f_X(a-y)dy.$$

Now the integrand is 0 unless $0 \leq a - y \leq 1$ (i.e. unless $a - 1 \leq y \leq a$) and then it is 1. So if $0 \leq a \leq 1$ 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,

$$f_{X+Y}(a) = \begin{cases} a & 0 \leq a \leq 1 \\ 2 - a & 1 < a < 2 \\ 0 & \text{otherwise} \blacksquare \end{cases}$$

Example 31.7

Let X and Y be two independent exponential random variables with common parameter λ . Compute $f_{X+Y}(a)$.

Solution.

We have

$$f_X(a) = f_Y(a) = \begin{cases} \lambda e^{-\lambda a} & 0 \leq a \\ 0 & \text{otherwise} \end{cases}$$

If $a \geq 0$ 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,

$$f_{X+Y}(a) = \begin{cases} a\lambda^2 e^{-\lambda a} & 0 \leq a \\ 0 & \text{otherwise} \blacksquare \end{cases}$$

Example 31.8

Let X and Y be two independent random variables, each with the standard normal density. Compute $f_{X+Y}(a)$.

Solution.

We have

$$f_X(a) = f_Y(a) = \frac{1}{\sqrt{2\pi}} e^{-\frac{a^2}{2}}.$$

By Theorem 31.1 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 31.9

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 31.1 we have

$$\begin{aligned} f_{X+Y}(a) &= \frac{1}{\Gamma(s)\Gamma(t)} \int_0^a \lambda e^{-\lambda(a-y)} [\lambda(a-y)]^{s-1} \lambda e^{-\lambda y} (\lambda y)^{t-1} dy \\ &= \frac{\lambda^{s+t} e^{-\lambda a}}{\Gamma(s)\Gamma(t)} \int_0^a (a-y)^{s-1} y^{t-1} dy \\ &= \frac{\lambda^{s+t} e^{-\lambda a} a^{s+t-1}}{\Gamma(s)\Gamma(t)} \int_0^1 (1-x)^{s-1} x^{t-1} dx, \quad x = \frac{y}{a} \end{aligned}$$

But we have from Theorem 29.8 that

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

Thus,

$$f_{X+Y}(a) = \frac{\lambda e^{-\lambda a} (\lambda a)^{s+t-1}}{\Gamma(s+t)} \blacksquare$$

Example 31.10

The percentages of copper and iron in a certain kind of ore are, respectively, X and Y . If the joint density of these two random variables is given by

$$f_{XY}(x, y) = \begin{cases} \frac{3}{11}(5x + y) & x, y > 0, \quad x + 2y < 2 \\ 0 & \text{elsewhere} \end{cases}$$

use the distribution function technique to 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 the figure we see that $F(a) = 0$ if $a < 0$. Now, the two lines $x + y = a$ and $x + 2y = 2$ intersect at $(2a - 2, 2 - a)$.

If $0 \leq a < 1$ then

$$F_Z(a) = P(Z \leq a) = \int_0^a \int_0^{a-y} \frac{3}{11}(5x + y) dx dy = \frac{3}{11} a^3$$

If $1 \leq a < 2$ then

$$\begin{aligned} F_Z(a) &= P(Z \leq a) \\ &= \int_0^{2-a} \int_0^{a-y} \frac{3}{11}(5x + y) dx dy + \int_{2-a}^1 \int_0^{2-2y} \frac{3}{11}(5x + y) dx dy \\ &= \frac{3}{11} \left(-\frac{7}{3} a^3 + 9a^2 - 8a + \frac{7}{3} \right) \end{aligned}$$

If $a \geq 2$ then $F_Z(a) = 1$. Differentiating with respect to a we find

$$f_Z(a) = \begin{cases} \frac{9}{11}a^2 & 0 < a \leq 1 \\ \frac{3}{11}(-7a^2 + 18a - 8) & 1 < a < 2 \\ 0 & \text{elsewhere} \blacksquare \end{cases}$$

Problems

Problem 31.13

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 31.14

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 31.15

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 31.16

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 31.17

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 31.18

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 31.19

A device containing two key components fails when and only when both components fail. The lifetime, T_1 and T_2 , of these components are independent with a common density function given by

$$f_{T_1}(t) = f_{T_2}(t) = \begin{cases} e^{-t} & t > 0 \\ 0 & \text{otherwise} \end{cases}$$

The cost, X , of operating the device until failure is $2T_1 + T_2$. Find the density function of X .

Problem 31.20

Let X and Y be independent random variables with density functions

$$f_X(x) = \begin{cases} \frac{1}{2} & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{2} & 3 \leq y \leq 5 \\ 0 & \text{otherwise} \end{cases}$$

Find the probability density function of $X + Y$.

Problem 31.21

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 31.22

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 31.23

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 31.24

The life (in days) of a certain machine has an exponential distribution with a mean of 1 day. The machine comes supplied with one spare. Find the density function (t measure in days) of the combined life of the machine and its spare if the life of the spare has the same distribution as the first machine, but is independent of the first machine.

Problem 31.25

X_1 and X_2 are independent exponential random variables each with a mean of 1. Find $P(X_1 + X_2 < 1)$.

32 Conditional Distributions: Discrete Case

Recall that for any two events E and F the conditional probability of E given F is defined by

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

provided that $P(F) > 0$.

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

$$\begin{aligned} p_{X|Y}(x|y) &= P(X = x|Y = y) \\ &= \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= \frac{p_{XY}(x, y)}{p_Y(y)} \end{aligned} \tag{32.1}$$

provided that $p_Y(y) > 0$.

Example 32.1

Suppose you and me are tossing two fair coins independently, and we will stop as soon as each one of us gets a head.

- Find the chance that we stop simultaneously.
- Find the conditional distribution of the number of coin tosses given that we stop simultaneously.

Solution.

(a) Let X be the number of times I have to toss my coin before getting a head, and Y be the number of times you have to toss your coin before getting a head. So X and Y are independent identically distributed geometric random variables with parameter $p = \frac{1}{2}$. Thus,

$$\begin{aligned} P(X = Y) &= \sum_{k=1}^{\infty} P(X = k, Y = k) = \sum_{k=1}^{\infty} P(X = k)P(Y = k) \\ &= \sum_{k=1}^{\infty} \frac{1}{4^k} = \frac{1}{3} \end{aligned}$$

(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

$$P(X = k|Y = k) = \frac{P(X = k, Y = k)}{P(X = Y)} = \frac{\frac{1}{4^k}}{\frac{1}{3}} = \frac{3}{4} \left(\frac{1}{4}\right)^{k-1}.$$

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 (32.1). We also mention one more use of (32.1):

$$\begin{aligned} p_X(x) &= \sum_y p_{XY}(x, y) \\ &= \sum_y p_{X|Y}(x|y)p_Y(y) \end{aligned} \quad (32.2)$$

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 (32.2).

The conditional cumulative distribution of X given that $Y = y$ is defined by

$$\begin{aligned} F_{X|Y}(x|y) &= P(X \leq x|Y = y) \\ &= \sum_{a \leq x} p_{X|Y}(a|y) \end{aligned}$$

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 32.1

If X and Y are independent, then

$$p_{X|Y}(x|y) = p_X(x).$$

Proof.

We have

$$\begin{aligned} p_{X|Y}(x|y) &= P(X = x|Y = y) \\ &= \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= \frac{P(X = x)P(Y = y)}{P(Y = y)} = P(X = x) = p_X(x) \blacksquare \end{aligned}$$

Example 32.2

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 32.3

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

$$\begin{aligned}
 P(X = k | X + Y = n) &= \frac{P(X = k, X + Y = n)}{P(X + Y = n)} \\
 &= \frac{P(X = k, Y = n - k)}{P(X + Y = n)} \\
 &= \frac{P(X = k)P(Y = n - k)}{P(X + Y = n)} \\
 &= \frac{e^{-\lambda_1} \lambda_1^k e^{-\lambda_2} \lambda_2^{n-k}}{k! (n-k)!} \left[\frac{e^{-(\lambda_1 + \lambda_2)} (\lambda_1 + \lambda_2)^n}{n!} \right]^{-1} \\
 &= \frac{n!}{k!(n-k)!} \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^n} \\
 &= \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-k}
 \end{aligned}$$

In other words, the conditional mass distribution function of X given that $X + Y = n$, is the binomial distribution with parameters n and $\frac{\lambda_1}{\lambda_1 + \lambda_2}$ ■

Problems

Problem 32.1

Given the following table.

$X \setminus Y$	$Y=0$	$Y=1$	$p_X(x)$
$X=0$.4	.1	.5
$X=1$.2	.3	.5
$p_Y(y)$.6	.4	1

Find $p_{X|Y}(x|y)$ where $Y = 1$.

Problem 32.2

Choose a number X from the set $\{1, 2, 3, 4, 5\}$ and then choose a number Y from $\{1, 2, \dots, X\}$.

- Find the joint mass function of X and Y .
- Find the conditional mass function of X given that $Y = i$.
- Are X and Y independent?

Problem 32.3

Consider the following hypothetical joint distribution of X , a person's grade on the AP calculus exam (a number between 1 and 5), and their grade Y in their high school calculus course, which we assume was $A = 4$, $B = 3$, or $C = 2$.

$X \setminus Y$	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 $P(X|Y = 4)$ for $X = 3, 4, 5$.

Problem 32.4

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

$X \setminus Y$	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 32.5

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 $x = 1, 2, \dots, 6$. Are X and Y independent?

Problem 32.6

Let X and Y be discrete random variables with joint probability function

$$p_{XY}(x, y) = \begin{cases} \frac{n!y^x(pe^{-1})^y(1-p)^{n-y}}{y!(n-y)!x!} & y = 0, 1, \dots, n; x = 0, 1, \dots \\ 0 & \text{otherwise} \end{cases}$$

(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 32.7

Let X and Y have the joint probability function $p_{XY}(x, y)$ described as follows:

$X \setminus Y$	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 32.8

Let X and Y be random variables with joint probability mass function

$$p_{XY}(x, y) = c(1 - 2^{-x})^y$$

where $x = 0, 1, \dots, N - 1$ and $y = 0, 1, 2, \dots$

- (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 32.9

Let X and Y be identically independent Poisson random variables with parameter λ . Find $P(X = k|X + Y = n)$.

Problem 32.10

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 $Y = 1$.

33 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

$$P(a < X < b|Y = y) = \int_a^b f_{X|Y}(x|y)dx.$$

Then for δ very small we have (See Remark 22.1)

$$P(x \leq X \leq x + \delta|Y = y) \approx \delta f_{X|Y}(x|y).$$

On the other hand, for small ϵ we have

$$\begin{aligned} P(x \leq X \leq x + \delta|Y = y) &\approx P(x \leq X \leq x + \delta|y \leq Y \leq y + \epsilon) \\ &= \frac{P(x \leq X \leq x + \delta, y \leq Y \leq y + \epsilon)}{P(y \leq Y \leq y + \epsilon)} \\ &\approx \frac{\delta \epsilon f_{XY}(x, y)}{\epsilon f_Y(y)} \end{aligned}$$

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 33.1

Suppose X and Y have the following joint density

$$f_{XY}(x, y) = \begin{cases} \frac{1}{2} & |X| + |Y| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- (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 $|x| \leq 1$. 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 $(-\frac{1}{2}, \frac{1}{2})$ ■

Example 33.2

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 $P(Y \geq \frac{1}{2})$.

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 33.1 we find

$$\begin{aligned}
 P(Y \geq \frac{1}{2}) &= \int_0^{\frac{1}{2}} \int_{1-x}^1 \frac{1}{x} dy dx + \int_{\frac{1}{2}}^1 \int_{\frac{1}{2}}^1 \frac{1}{x} dy dx \\
 &= \int_0^{\frac{1}{2}} \frac{1 - (1-x)}{x} dx + \int_{\frac{1}{2}}^1 \frac{1/2}{x} dx \\
 &= \frac{1 + \ln 2}{2} \blacksquare
 \end{aligned}$$

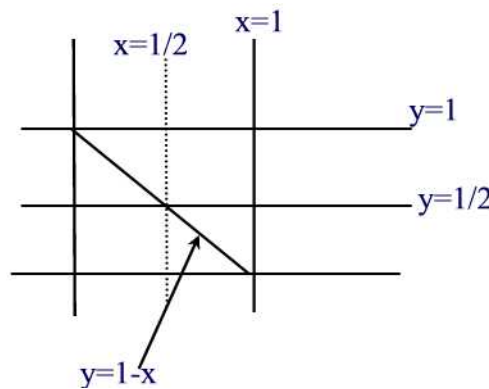


Figure 33.1

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

$$F_{X|Y}(x|y) = P(X \leq x | Y = y) = \int_{-\infty}^x f_{X|Y}(t|y) dt.$$

From this definition, it follows

$$f_{X|Y}(x|y) = \frac{\partial}{\partial x} F_{X|Y}(x|y).$$

Example 33.3

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,

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{XY}(x,y)}{f_Y(y)} \\ &= \frac{x(2-x-y)}{\frac{2}{3} - \frac{y}{2}} \\ &= \frac{6x(2-x-y)}{4-3y} \quad \blacksquare \end{aligned}$$

Example 33.4

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 $P(X > 1|Y = y)$.

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,

$$\begin{aligned} P(X > 1|Y = y) &= \int_1^{\infty} \frac{1}{y} e^{-\frac{x}{y}} dx \\ &= -e^{-\frac{x}{y}} \Big|_1^{\infty} = e^{-\frac{1}{y}} \blacksquare \end{aligned}$$

We end this section with the following theorem.

Theorem 33.1

Continuous random variables X and Y 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 33.5

Let X and Y be two continuous random variables with joint density function

$$f_{XY}(x, y) = \begin{cases} c & 0 \leq y < x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Find $f_X(x)$, $f_Y(y)$ and $f_{X|Y}(x|1)$.
 (b) Are X and Y independent?

Solution.

(a) We have

$$\begin{aligned} f_X(x) &= \int_0^x c dy = cx, \quad 0 \leq x \leq 2 \\ f_Y(y) &= \int_y^2 c dx = c(2 - y), \quad 0 \leq y \leq 2 \end{aligned}$$

and

$$f_{X|Y}(x|1) = \frac{f_{XY}(x, 1)}{f_Y(1)} = \frac{c}{c} = 1, \quad 0 \leq x \leq 1.$$

- (b) Since $f_{X|Y}(x|1) \neq f_X(x)$ then X and Y are dependent \blacksquare

Problems

Problem 33.1

Let X and Y be two random variables with joint density function

$$f_{XY}(x, y) = \begin{cases} 5x^2y & -1 \leq x \leq 1, y \leq |x| \\ 0 & \text{otherwise} \end{cases}$$

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 33.2

Suppose that X and Y have joint distribution

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

Find $f_{X|Y}(x|y)$, the conditional probability density function of X given $Y = y$.

Problem 33.3

Suppose that X and Y have joint distribution

$$f_{XY}(x, y) = \begin{cases} \frac{3y^2}{x^3} & 0 \leq y < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $f_{Y|X}(x|y)$, the conditional probability density function of Y given $X = x$.

Problem 33.4

The joint density function of X and Y is given by

$$f_{XY}(x, y) = \begin{cases} xe^{-x(y+1)} & x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find the conditional density of X given $Y = y$ and that of Y given $X = x$.

Problem 33.5

Let X and Y be continuous random variables with conditional and marginal p.d.f.'s given by

$$f_X(x) = \frac{x^3 e^{-x}}{6} I_{(0, \infty)}(x)$$

and

$$f_{Y|X}(x|y) = \frac{3y^2}{x^3} I_{(0,x)}.$$

- (a) Find the joint p.d.f. of X and Y .
 (b) Find the conditional p.d.f. of X given $Y = y$.

Problem 33.6

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 $P(Y < \frac{1}{2} | X > \frac{1}{2})$.

Problem 33.7

The joint probability density function of the random variables X and Y is given by

$$f_{XY}(x, y) = \begin{cases} \frac{1}{3}x - y + 1 & 1 \leq x \leq 2, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Find the conditional probability density function of X given $Y = y$.
 (b) Find $P(X < \frac{3}{2} | Y = \frac{1}{2})$.

Problem 33.8 †

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 $P(Y < X | X = \frac{1}{3})$.

Problem 33.9 †

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 33.10 ‡

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 33.11 ‡

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 0.500?

Problem 33.12 ‡

You are given the following information about N , the annual number of claims for a randomly selected insured:

$$P(N = 0) = \frac{1}{2}$$

$$P(N = 1) = \frac{1}{3}$$

$$P(N > 1) = \frac{1}{6}$$

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 $P(4 < S < 8)$.

Problem 33.13

Let Y have a uniform distribution on the interval $(0, 1)$, and let the con-

ditional 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 33.14

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 .

34 Joint Probability Distributions of Functions of Random Variables

Theorem 28.1 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 34.1

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 recall 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.

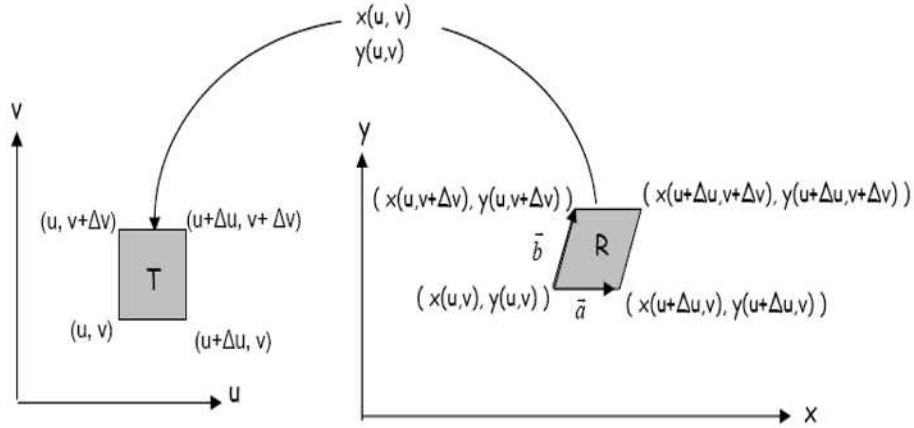


Figure 34.1

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 34.1. 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 are

$$\vec{a} = [x(u + \Delta u, v) - x(u, v)]\vec{i} + [y(u + \Delta u, v) - y(u, v)]\vec{j} \approx \frac{\partial x}{\partial u} \Delta u \vec{i} + \frac{\partial y}{\partial u} \Delta u \vec{j}$$

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

$$\begin{aligned} \int_R f(x, y) dx dy &\approx \sum_{j=1}^m \sum_{i=1}^n f(x_{ij}, y_{ij}) \cdot \text{Area of } R_{ij} \\ &\approx \sum_{j=1}^m \sum_{i=1}^n f(u_{ij}, v_{ij}) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| \Delta u \Delta v \end{aligned}$$

where (x_{ij}, y_{ij}) in R_{ij} corresponds to a point (u_{ij}, v_{ij}) in T_{ij} . Now, letting $m, n \rightarrow \infty$ to obtain

$$\int_R f(x, y) dx dy = \int_T f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv.$$

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

$$\begin{aligned} P((X, Y) \in R) &= \int \int_R f_{XY}(x, y) dx dy \\ &= \int \int_T f_{XY}(x(u, v), y(u, v)) |J(x(u, v), y(u, v))|^{-1} du dv \\ &= P((U, V) \in T) \blacksquare \end{aligned}$$

Example 34.1

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 34.2

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 $J(x, y) = -2$ then

$$f_{UV}(u, v) = \frac{1}{4\pi}e^{-\frac{(\frac{u+v}{2})^2 + (\frac{u-v}{2})^2}{2}} = \frac{1}{4\pi}e^{-\frac{u^2+v^2}{4}} \blacksquare$$

Example 34.3

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 $V = XY$.

- Find the joint density function of U and V .
- Find the marginal density of U and V .
- 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 34.1, we find

$$f_{UV}(u, v) = \frac{1}{2u}f_{XY}\left(\sqrt{uv}, \sqrt{\frac{v}{u}}\right) = \frac{2v}{u}, \quad 0 < uv < 1, \quad 0 < \frac{v}{u} < 1$$

and 0 otherwise. The region where f_{UV} is shown in Figure 34.2

(b) The marginal density of U is

$$f_U(u) = \int_0^u \frac{2v}{u} dv = v, \quad u \leq 1$$

$$f_U(u) = \int_0^{\frac{1}{u}} \frac{2v}{u} dv = \frac{1}{u^3}, \quad u > 1$$

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 $f_{UV}(u, v) \neq f_U(u)f_V(v)$ then U and V are dependent ■

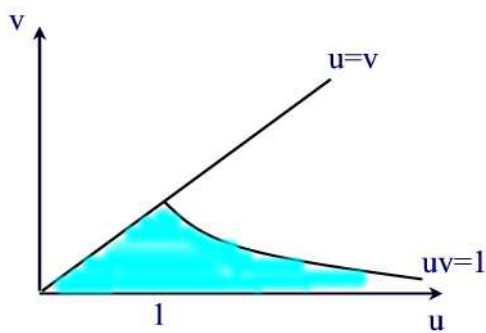


Figure 34.2

Problems

Problem 34.1

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 34.2

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 34.3

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 34.4

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 34.5

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 34.6

Let X_1 and X_2 be two continuous random variables with joint density function

$$f_{X_1 X_2}(x_1, x_2) = \begin{cases} e^{-(x_1+x_2)} & x_1 \geq 0, x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

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 34.7

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 $f_{Y_1 Y_2}(y_1, y_2)$.

Problem 34.8

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 34.9

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 34.10

Let X and Y be two random variables with joint density function f_{XY} . Compute the pdf of $U = Y - X$.

Problem 34.11

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 34.12

The daily amounts of Coke and Diet Coke sold by a vendor are independent and follow exponential distributions with a mean of 1 (units are 100s of gallons). Use Theorem 34.1 to obtain the distribution of the ratio of Coke to Diet Coke sold.

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.

35 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 $g : S_X \times S_Y \rightarrow \mathbb{R}$ 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

$$E(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{XY}(x, y) dx dy$$

if X and Y are continuous with joint probability density function $f_{XY}(x, y)$.

Example 35.1

Let X and Y be two discrete random variables with joint probability mass function:

$$p_{XY}(1, 1) = \frac{1}{3}, p_{XY}(1, 2) = \frac{1}{8}, p_{XY}(2, 1) = \frac{1}{2}, p_{XY}(2, 2) = \frac{1}{24}$$

Find the expected value of $g(X, Y) = XY$.

Solution.

The expected value of the function $g(X, Y) = XY$ is calculated as follows:

$$\begin{aligned} E(g(X, Y)) &= E(XY) = \sum_{x=1}^2 \sum_{y=1}^2 xyp_{XY}(x, y) \\ &= (1)(1)\left(\frac{1}{3}\right) + (1)(2)\left(\frac{1}{8}\right) + (2)(1)\left(\frac{1}{2}\right) + (2)(2)\left(\frac{1}{24}\right) \\ &= \frac{7}{4} \blacksquare \end{aligned}$$

An important application of the above definition is the following result.

Proposition 35.1

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 35.2

During commencement at the United States Naval academy, there is a tradition that everyone throws their hat into the air at the conclusion of the ceremony. Later that day, the hats are collected and each graduate is given a hat at random. Assuming that all of the hats are indistinguishable from the others (so hats are really given back at random) and that there are 1000 graduates, calculate $E(X)$, where X is the number of people who receive the hat that they wore during the ceremony.

Solution.

Let

$$X_i = \begin{cases} 0 & \text{if the } i\text{th person does not get his/her hat} \\ 1 & \text{if the } i\text{th person gets his/her hat} \end{cases}$$

Then $X = X_1 + X_2 + \cdots + X_{1000}$ and

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_{1000}).$$

But for $1 \leq i \leq 1000$

$$E(X_0) = 0 \times \frac{999}{1000} + 1 \times \frac{1}{1000} = \frac{1}{1000}.$$

Hence,

$$E(X) = 1000E(X_i) = 1 \blacksquare$$

Example 35.3 (Sample Mean)

Let X_1, X_2, \dots, 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

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

We call \bar{X} the **sample mean**. Find $E(\bar{X})$.

Solution.

The expected value of \bar{X} is

$$E(\bar{X}) = E\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \frac{1}{n} \sum_{i=1}^n E(X_i) = \mu.$$

Because of this result, when the distribution mean μ is unknown, the sample mean is often used in statistics to estimate it \blacksquare

The following property is known as the monotonicity property of the expected value.

Proposition 35.2

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 35.3 (*Boole's Inequality*)

For any events A_1, A_2, \dots, A_n we have

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

Proof.

For $i = 1, \dots, n$ 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

$$E(Y) = P\{\text{at least one of the } A_i \text{ occur}\} = P\left(\bigcup_{i=1}^n A_i\right).$$

Thus, the result follows. Note that for any set A we have

$$E(I_A) = \int I_A(x)f(x)dx = \int_A f(x)dx = P(A) \quad \blacksquare$$

Proposition 35.4

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 35.5

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

$$\begin{aligned} E(g(X)h(Y)) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)f_{XY}(x, y)dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dx dy \\ &= \left(\int_{-\infty}^{\infty} h(y)f_Y(y)dy \right) \left(\int_{-\infty}^{\infty} g(x)f_X(x)dx \right) \\ &= E(h(Y))E(g(X)) \blacksquare \end{aligned}$$

We next give a simple example to show that the expected values need not multiply if the random variables are not independent.

Example 35.4

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 35.5

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, $X_i Y_i = 0$ 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

$$XY = \sum_{1 \leq i \neq j \leq 10} X_i Y_j.$$

Hence,

$$E(XY) = \sum_{1 \leq i \neq j \leq 10} \sum_{1 \leq i \neq j \leq 10} E(X_i Y_j) = \sum_{1 \leq i \neq j \leq 10} E(X_i) E(Y_j) = 90 \times \frac{1}{3} \times \frac{1}{3} = 10.$$

(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 35.6 (*Markov's Inequality*)

If $X \geq 0$ and $c > 0$ then $P(X \geq c) \leq \frac{E(X)}{c}$.

Proof.

Let $c > 0$. Define

$$I = \begin{cases} 1 & \text{if } X \geq c \\ 0 & \text{otherwise} \end{cases}$$

Since $X \geq 0$ then $I \leq \frac{X}{c}$. Taking expectations of both side we find $E(I) \leq \frac{E(X)}{c}$. Now the result follows since $E(I) = P(X \geq c)$ ■

Example 35.6

Let X be a non-negative random variable. Let a be a positive constant. Prove that $P(X \geq a) \leq \frac{E(e^{tX})}{e^{ta}}$ for all $t \geq 0$.

Solution.

Applying Markov's inequality we find

$$P(X \geq a) = P(tX \geq ta) = P(e^{tX} \geq e^{ta}) \leq \frac{E(e^{tX})}{e^{ta}} \blacksquare$$

As an important application of the previous result we have

Proposition 35.7

If $X \geq 0$ and $E(X) = 0$ then $P(X = 0) = 1$.

Proof.

Since $E(X) = 0$ then by the previous result $P(X \geq c) = 0$ for all $c > 0$. But

$$P(X > 0) = P\left(\bigcup_{n=1}^{\infty} \left(X > \frac{1}{n}\right)\right) \leq \sum_{n=1}^{\infty} P\left(X > \frac{1}{n}\right) = 0.$$

Hence, $P(X > 0) = 0$. Since $X \geq 0$ then $1 = P(X \geq 0) = P(X = 0) + P(X > 0) = P(X = 0) \blacksquare$

Corollary 35.1

Let X be a random variable. If $Var(X) = 0$, then $P(X = E(X)) = 1$.

Proof.

Suppose that $Var(X) = 0$. Since $(X - E(X))^2 \geq 0$ and $Var(X) = E((X - E(X))^2)$ then by the previous result we have $P(X - E(X) = 0) = 1$. That is, $P(X = E(X)) = 1 \blacksquare$

Example 35.7 (*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 i th trial. Then $E(X_i) = 0(1-p) + 1p = p$. Since $X = X_1 + X_2 + \cdots + X_n$ then $E(X) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n p = np \blacksquare$

Problems

Problem 35.1

Let X and Y be independent random variables, both being equally likely to be any of the numbers $1, 2, \dots, m$. Find $E(|X - Y|)$.

Problem 35.2

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 35.3

Let X and Y be two independent uniformly distributed random variables in $[0, 1]$. Find $E(|X - Y|)$.

Problem 35.4

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 35.5

Suppose that $E(X) = 5$ and $E(Y) = -2$. Find $E(3X + 4Y - 7)$.

Problem 35.6

Suppose that X and Y are independent, and that $E(X) = 5$, $E(Y) = -2$. Find $E[(3X - 4)(2Y + 7)]$.

Problem 35.7

An accident occurs at a point X that is uniformly distributed on a road of length L . At the time of the accident an ambulance is at location Y that is also uniformly distributed on the road. Assuming that X and Y are independent, find the expected distance between the ambulance and the point of the accident.

Problem 35.8

A group of N people throw their hats into the center of a room. The hats are mixed, and each person randomly selects one. Find the expected number of people that select their own hat.

Problem 35.9

Twenty people, consisting of 10 married couples, are to be seated at five different tables, with four people at each table. If the seating is done at random, what is the expected number of married couples that are seated at the same table?

Problem 35.10

Suppose that A and B each randomly, and independently, choose 3 out of 10 objects. Find the expected number of objects

- (a) chosen by both A and B .
- (b) not chosen by either A or B
- (c) chosen exactly by one of A and B .

Problem 35.11

If $E(X) = 1$ and $\text{Var}(X) = 5$ find

- (a) $E[(2 + X)^2]$
- (b) $\text{Var}(4 + 3X)$

Problem 35.12 ‡

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 35.13 ‡

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 $0 \leq t_1 \leq t_2 \leq L$, where L is a positive constant.

Determine the expected value of the sum of the squares of T_1 and T_2 .

Problem 35.14

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 $E[(X + 1)^2(Y - 1)^2]$.

36 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

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))].$$

An alternative expression that is sometimes more convenient is

$$\begin{aligned} \text{Cov}(X, Y) &= E(XY - E(X)Y - XE(Y) + E(X)E(Y)) \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

Recall that for independent X, Y we have $E(XY) = E(X)E(Y)$ and so $\text{Cov}(X, Y) = 0$. 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

$$P(X = 0) = P(X = 1) = P(X = -1) = \frac{1}{3}$$

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

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0.$$

Useful facts are collected in the next result.

Theorem 36.1

- (a) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ (Symmetry)
- (b) $\text{Cov}(X, X) = \text{Var}(X)$
- (c) $\text{Cov}(aX, Y) = a\text{Cov}(X, Y)$
- (d) $\text{Cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m \text{Cov}(X_i, Y_j)$

Proof.

- (a) $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = E(YX) - E(Y)E(X) = \text{Cov}(Y, X)$.
- (b) $\text{Cov}(X, X) = E(X^2) - (E(X))^2 = \text{Var}(X)$.
- (c) $\text{Cov}(aX, Y) = E(aXY) - E(aX)E(Y) = aE(XY) - aE(X)E(Y) = a(E(XY) - E(X)E(Y)) = a\text{Cov}(X, Y)$.
- (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

$$\begin{aligned} \text{Cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) &= E\left[\left(\sum_{i=1}^n X_i - \sum_{i=1}^n E(X_i)\right)\left(\sum_{j=1}^m Y_j - \sum_{j=1}^m E(Y_j)\right)\right] \\ &= E\left[\sum_{i=1}^n (X_i - E(X_i)) \sum_{j=1}^m (Y_j - E(Y_j))\right] \\ &= E\left[\sum_{i=1}^n \sum_{j=1}^m (X_i - E(X_i))(Y_j - E(Y_j))\right] \\ &= \sum_{i=1}^n \sum_{j=1}^m E[(X_i - E(X_i))(Y_j - E(Y_j))] \\ &= \sum_{i=1}^n \sum_{j=1}^m \text{Cov}(X_i, Y_j) \blacksquare \end{aligned}$$

Example 36.1

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,

$$\text{Cov}(X + Y, X + 1.2Y) = E((X + Y)(X + 1.2Y)) - E(X + Y)E(X + 1.2Y)$$

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,

$$\text{Cov}(X + Y, X + 1.2Y) = 2.2E(XY) + 89.08 - 160.8 = 2.2E(XY) - 71.72$$

To complete the calculation, it remains to find $E(XY)$. To this end we make use of the still unused relation $\text{Var}(X + Y) = 8$

$$\begin{aligned} 8 &= \text{Var}(X + Y) = E((X + Y)^2) - (E(X + Y))^2 \\ &= E(X^2) + 2E(XY) + E(Y^2) - (E(X) + E(Y))^2 \\ &= 27.4 + 2E(XY) + 51.4 - (5 + 7)^2 = 2E(XY) - 65.2 \end{aligned}$$

so $E(XY) = 36.6$. Substituting this above gives $\text{Cov}(X + Y, X + 1.2Y) = (2.2)(36.6) - 71.72 = 8.8$ ■

Example 36.2

On reviewing data on smoking (X , number of packages of cigarettes smoked per year) income (Y , in thousands per year) and health (Z number of visits to the family physician per year) for a sample of males it is found that $E(X) = 10$, $\text{Var}(X) = 25$, $E(Y) = 50$, $\text{Var}(Y) = 100$, $E(Z) = 6$, $\text{Var}(Z) = 4$, $\text{Cov}(X, Y) = 10$, and $\text{Cov}(X, Z) = 3.5$. Dr. I.P. Freely, a young statistician, attempts to describe the variable Z in terms of X and Y by the relation $Z = X + cY$, where c is a constant to be determined. Dr. Freely methodology for determining c is to find the value of c for which $\text{Cov}(Y, Z) = 3.5$ when Z is replaced by $X + cY$. What value of c does Dr. Freely find?

Solution.

We have

$$\begin{aligned} \text{Cov}(X, Z) &= \text{Cov}(X, X + cY) = \text{Cov}(X, X) + c\text{Cov}(X, Y) \\ &= \text{Var}(X) + c\text{Cov}(X, Y) = 25 + c(-10) = 3.5 \end{aligned}$$

Solving for c we find $c = 2.15$ ■

Using (b) and (d) in the previous theorem with $Y_j = X_j$, $j = 1, 2, \dots, n$ we find

$$\begin{aligned} \text{Var} \left(\sum_{i=1}^n X_i \right) &= \text{Cov} \left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i \right) \\ &= \sum_{i=1}^n \sum_{i=1}^n \text{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \sum_{i=1}^n \text{Cov}(X_i, X_j) \end{aligned}$$

Since each pair of indices $i \neq j$ appears twice in the double summation, the above reduces to

$$\text{Var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i < j} \sum_{i=1}^n \text{Cov}(X_i, X_j).$$

In particular, if X_1, X_2, \dots, X_n are pairwise independent then

$$\text{Var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{Var}(X_i).$$

Example 36.3

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

$$\begin{aligned} \text{Var}(Z) &= \text{Var}(3X - Y - 5) = \text{Var}(3X - Y) \\ &= \text{Var}(3X) + \text{Var}(-Y) = 9\text{Var}(X) + \text{Var}(Y) = 11 \quad \blacksquare \end{aligned}$$

Example 36.4

An insurance policy pays a total medical benefit consisting of a part paid to the surgeon, X , and a part paid to the hospital, Y , so that the total

benefit 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 benefit after these increases?

Solution.

We need to compute $\text{Var}(X + 100 + 1.1Y)$. Since adding constants does not change the variance, this is the same as $\text{Var}(X + 1.1Y)$, which expands as follows:

$$\begin{aligned}\text{Var}(X + 1.1Y) &= \text{Var}(X) + \text{Var}(1.1Y) + 2\text{Cov}(X, 1.1Y) \\ &= \text{Var}(X) + 1.21\text{Var}(Y) + 2(1.1)\text{Cov}(X, Y)\end{aligned}$$

We are given that $\text{Var}(X) = 5,000$, $\text{Var}(Y) = 10,000$, so the only remaining unknown quantity is $\text{Cov}(X, Y)$, which can be computed via the general formula for $\text{Var}(X + Y)$:

$$\begin{aligned}\text{Cov}(X, Y) &= \frac{1}{2}(\text{Var}(X + Y) - \text{Var}(X) - \text{Var}(Y)) \\ &= \frac{1}{2}(17,000 - 5,000 - 10,000) = 1,000\end{aligned}$$

Substituting this into the above formula, we get the answer:

$$\text{Var}(X + 1.1Y) = 5,000 + 1.21(10,000) + 2(1.1)(1,000) = 19,520 \blacksquare$$

Example 36.5

Let \bar{X} be the sample mean of n independent random variables X_1, X_2, \dots, X_n . Find $\text{Var}(\bar{X})$.

Solution.

By independence we have

$$\text{Var}(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) \blacksquare$$

The following result is known as the Cauchy Schwartz inequality.

Theorem 36.2

Let X and Y be two random variables. Then

$$\text{Cov}(X, Y)^2 \leq \text{Var}(X)\text{Var}(Y)$$

with equality if and only if X and Y are linearly related, i.e.,

$$Y = aX + b$$

for some constants a and b with $a \neq 0$.

Proof.

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

$$\rho(X, Y) = \frac{E(aX + bX^2) - E(X)E(a + bX)}{\sqrt{Var(X)b^2Var(X)}} = \frac{bVar(X)}{|b|Var(X)} = \text{sign}(b).$$

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 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 36.1 shows some examples of data pairs and their correlation.

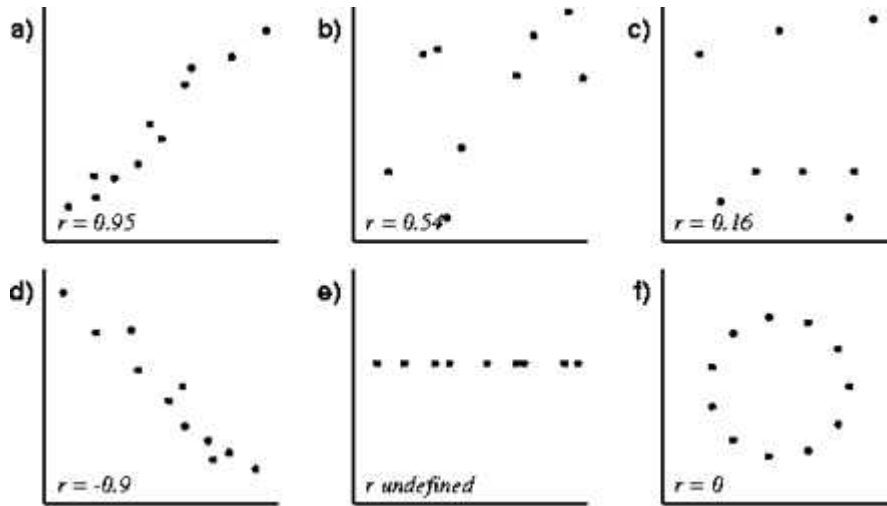


Figure 36.1

Problems

Problem 36.1

If X and Y are independent and identically distributed with mean μ and variance σ^2 , find $E[(X - Y)^2]$.

Problem 36.2

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 36.3

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 36.4

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.

- Find the joint pdf of X and Y .
- Find the covariance and the correlation of X and Y .

Problem 36.5

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 36.6

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

- $X_1 + X_2$ and $X_2 + X_3$.
- $X_1 + X_2$ and $X_3 + X_4$.

Problem 36.7

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 $\text{Cov}(X, Y)$.

Problem 36.8

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 36.9

Let X and Y be continuous random variables with joint pdf

$$f_{XY}(x, y) = \begin{cases} 3x & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $\text{Cov}(X, Y)$ and $\rho(X, Y)$.

Problem 36.10

Suppose that X and Y are random variables with $\text{Cov}(X, Y) = 3$. Find $\text{Cov}(2X - 5, 4Y + 2)$.

Problem 36.11 †

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 36.12 †

The profit for a new product is given by $Z = 3X - Y - 5$. X and Y are independent random variables with $\text{Var}(X) = 1$ and $\text{Var}(Y) = 2$.

What is the variance of Z ?

Problem 36.13 ‡

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 36.14 ‡

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 36.15 ‡

Let X and Y be continuous random variables with joint density function

$$f_{XY}(x, y) = \begin{cases} \frac{8}{3}xy & 0 \leq x \leq 1, x \leq y \leq 2x \\ 0 & \text{otherwise} \end{cases}$$

Find $\text{Cov}(X, Y)$

Problem 36.16 ‡

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 $\text{Cov}(X, Y)$ according to this model.

Problem 36.17 ‡

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 $\text{Var}(X + Y) = 8$.

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 $\text{Cov}(C_1, C_2)$.

Problem 36.18 ‡

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 36.19

Let X and Y be two independent random variables with densities

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{2} & 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

- Write down the joint pdf $f_{XY}(x, y)$.
- Let $Z = X + Y$. Find the pdf $f_Z(z)$. Simplify as much as possible.
- Find the expectation $E(X)$ and variance $\text{Var}(X)$. Repeat for Y .
- Compute the expectation $E(Z)$ and the variance $\text{Var}(Z)$.

Problem 36.20

Let X and Y be two random variables with joint pdf

$$f_{XY}(x, y) = \begin{cases} \frac{1}{2} & x > 0, y > 0, x + y < 2 \\ 0 & \text{otherwise} \end{cases}$$

- Let $Z = X + Y$. Find the pdf of Z .
- Find the pdf of X and that of Y .
- Find the expectation and variance of X .
- Find the covariance $\text{Cov}(X, Y)$.

Problem 36.21

Let X and Y be discrete random variables with joint distribution defined by the following table

$Y \setminus X$	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 $\text{Cov}(X, Y)$.

Problem 36.22

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 36.23

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 36.24

Let X and Y be discrete random variables with joint probability function $p_{XY}(x, y)$ given by the following table:

$X \setminus Y$	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 36.25

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 36.26

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 36.27

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 36.28

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 36.29

Given n independent random variables X_1, X_2, \dots, X_n each having the same variance σ^2 . Define $U = 2X_1 + X_2 + \dots + X_{n-1}$ and $V = X_2 + X_3 + \dots + X_{n-1} + 2X_n$. Find $\rho(U, V)$.

Problem 36.30

The following table gives the joint probability distribution for the numbers of washers (X) and dryers (Y) sold by an appliance store salesperson in a day.

$X \setminus Y$	0	1	2
0	0.25	0.08	0.05
1	0.12	0.20	0.10
2	0.03	0.07	0.10

- Give the marginal distributions of numbers of washers and dryers sold per day.
- Give the Expected numbers of Washers and Dryers sold in a day.
- Give the covariance between the number of washers and dryers sold per day.
- If the salesperson makes a commission of \$100 per washer and \$75 per dryer, give the average daily commission.

Problem 36.31

In a large class, on exam 1, the mean and standard deviation of scores were 72 and 15, respectively. For exam 2, the mean and standard deviation were 68 and 20, respectively. The covariance of the exam scores was 120. Give the mean and standard deviation of the sum of the two exam scores. Assume all students took both exams

37 Conditional Expectation

Since conditional probability measures are probability 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

$$\begin{aligned} E(X|Y = y) &= \sum_x xP(X = x|Y = y) \\ &= \sum_x xp_{X|Y}(x|y) \end{aligned}$$

where $p_{X|Y}$ is the conditional probability mass function of X , given that $Y = y$ which is given by

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{p(x, y)}{p_Y(y)}.$$

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 37.1

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

$$\begin{aligned}
 E(Y|X = 3) &= \sum_{y=1}^4 1y p_{Y|X}(y|3) \\
 &= \frac{p_{XY}(3, 1)}{p_X(3)} + \frac{2p_{XY}(3, 2)}{p_X(3)} + \frac{3p_{XY}(3, 3)}{p_X(3)} + \frac{4p_{XY}(3, 4)}{p_X(3)} \\
 &= 3 \cdot \frac{1}{3} + 4 \cdot \frac{2}{3} = \frac{11}{3} \blacksquare
 \end{aligned}$$

Example 37.2

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,

$$\begin{aligned} E(X|Y = y) &= \int_0^{\infty} \frac{x}{y} e^{-\frac{x}{y}} dx = - \left[x e^{-\frac{x}{y}} \Big|_0^{\infty} - \int_0^{\infty} e^{-\frac{x}{y}} dx \right] \\ &= - \left[x e^{-\frac{x}{y}} + y e^{-\frac{x}{y}} \right]_0^{\infty} = y \blacksquare \end{aligned}$$

Example 37.3

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 $f_X(x) = 0$, $x \leq 0$. 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 $x \geq 1$ we have

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy = \int_x^{\infty} f_{XY}(x, y) dy = \frac{\alpha-1}{\alpha} \frac{1}{x^\alpha}$$

- (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 $p_{X|Y}(x|y) = p(x)$ 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 37.1 (*Double Expectation Property*)

$$E(X) = E(E(X|Y))$$

Proof.

We give a proof in the case X and Y are continuous random variables.

$$\begin{aligned}
 E(E(X|Y)) &= \int_{-\infty}^{\infty} E(X|Y = y) f_Y(y) dy \\
 &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx \right) f_Y(y) dy \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y}(x|y) f_Y(y) dx dy \\
 &= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{XY}(x, y) dy dx \\
 &= \int_{-\infty}^{\infty} x f_X(x) dx = E(X) \blacksquare
 \end{aligned}$$

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

$$P(A) = E(X)$$

and for any random variable Y

$$E(X|Y = y) = P(A|Y = y).$$

Thus, by the double expectation property we have

$$\begin{aligned}
 P(A) = E(X) &= \sum_y E(X|Y = y) P(Y = y) \\
 &= \sum_y P(A|Y = y) p_Y(y)
 \end{aligned}$$

in the discrete case and

$$P(A) = \int_{-\infty}^{\infty} P(A|Y = y) f_Y(y) dy$$

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 37.1

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) $\text{Var}(X) = E[\text{Var}(X|Y)] + \text{Var}(E(X|Y))$.

Proof.

(a) We have

$$\begin{aligned} \text{Var}(X|Y) &= E[(X - E(X|Y))^2|Y = y] \\ &= E[(X^2 - 2XE(X|Y) + (E(X|Y))^2|Y] \\ &= E(X^2|Y) - 2E(X|Y)E(X|Y) + (E(X|Y))^2 \\ &= E(X^2|Y) - [E(X|Y)]^2 \end{aligned}$$

(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 $E(E(X|Y)) = E(X)$ then

$$\text{Var}(E(X|Y)) = E[(E(X|Y))^2] - (E(X))^2.$$

(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 37.2

$$\min_g E[(Y - g(X))^2] = E(Y - E(Y|X)).$$

Proof.

We have

$$\begin{aligned} E[(Y - g(X))^2] &= E[(Y - E(Y|X) + E(Y|X) - g(X))^2] \\ &= E[(Y - E(Y|X))^2] + E[(E(Y|X) - g(X))^2] \\ &\quad + 2E[(Y - E(Y|X))(E(Y|X) - g(X))] \end{aligned}$$

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 $g(X) = E(Y|X)$ ■

Problems

Problem 37.1

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 37.2

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 37.3

Let X and Y be independent exponentially distributed random variables with parameters μ and λ respectively. Using conditioning, find $P(X > Y)$.

Problem 37.4

Let X be uniformly distributed on $[0, 1]$. Find $E(X|X > 0.5)$.

Problem 37.5

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 37.6

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 37.7

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 37.8

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 37.9

Let X be an exponential random variable with $\lambda = 5$ and Y a uniformly distributed random variable on $(-3, X)$. Find $E(Y)$.

Problem 37.10

The number of people who pass by a store during lunch time (say from 12:00 to 1:00 pm) is a Poisson random variable with parameter $\lambda = 100$. Assume that each person may enter the store, independently of the other people, with a given probability $p = 0.15$. What is the expected number of people who enter the store during lunch time?

Problem 37.11

Let X and Y be discrete random variables with joint probability mass function defined by the following table

$X \setminus Y$	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 37.12 ‡

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:

$$P(X = 0, Y = 0) = 0.800$$

$$P(X = 1, Y = 0) = 0.050$$

$$P(X = 0, Y = 1) = 0.025$$

$$P(X = 1, Y = 1) = 0.125$$

Calculate $\text{Var}(Y|X = 1)$.

Problem 37.13 ‡

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 37.14 ‡

An actuary determines that the annual numbers of tornadoes in counties P and Q are jointly distributed as follows:

$X \setminus Y$	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 37.15

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 37.16

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 37.17

The number of stops X in a day for a delivery truck driver is Poisson with mean λ . Conditional on their being $X = x$ stops, the expected distance driven by the driver Y is Normal with a mean of αx miles, and a standard deviation of βx miles. Give the mean and variance of the numbers of miles she drives per day.

38 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 $p(x)$ we have

$$M_X(t) = \sum_x e^{tx} p(x)$$

and for a continuous random variable with pdf f ,

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx.$$

Example 38.1

Let X be a discrete random variable with pmf given by the following table

x	1	2	3	4	5
p(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 38.2

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 $n = 1, 2, \dots$. Our first result shows how to use the moment generating function to calculate moments.

Proposition 38.1

$$E(X^n) = M_X^n(0)$$

where

$$M_X^n(0) = \frac{d^n}{dt^n} M_X(t) \Big|_{t=0}$$

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

$$\begin{aligned} \frac{d}{dt} M_X(t) &= \frac{d}{dt} \int_{-\infty}^{\infty} e^{tx} f(x) dx = \int_{-\infty}^{\infty} \left(\frac{d}{dt} e^{tx} \right) f(x) dx \\ &= \int_{-\infty}^{\infty} x e^{tx} f(x) dx = E[X e^{tX}] \end{aligned}$$

Hence,

$$\frac{d}{dt} M_X(t) \Big|_{t=0} = E[X e^{tX}] \Big|_{t=0} = E(X).$$

By induction on n we find

$$\frac{d^n}{dt^n} M_X(t) \Big|_{t=0} = E[X^n e^{tX}] \Big|_{t=0} = E(X^n) \blacksquare$$

We next compute $M_X(t)$ for some common distributions.

Example 38.3

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

$$\begin{aligned} M_X(t) = E(e^{tX}) &= \sum_{k=0}^n e^{tk} C(n, k) p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n C(n, k) (pe^t)^k (1-p)^{n-k} = (pe^t + 1 - p)^n \end{aligned}$$

Differentiating yields

$$\frac{d}{dt}M_X(t) = npe^t(pe^t + 1 - p)^{n-1}$$

Thus

$$E(X) = \frac{d}{dt}M_X(t) \Big|_{t=0} = np.$$

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

$$\text{Var}(X) = E(X^2) - (E(X))^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p) \blacksquare$$

Example 38.4

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

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \lambda \blacksquare$$

Example 38.5

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

$$M_X'(t) = \frac{\lambda}{(\lambda-t)^2} \text{ and } M_X''(t) = \frac{2\lambda}{(\lambda-t)^3}.$$

Hence,

$$E(X) = M_X'(0) = \frac{1}{\lambda} \text{ and } E(X^2) = M_X''(0) = \frac{2}{\lambda^2}.$$

The variance of X is given by

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{1}{\lambda^2} \blacksquare$$

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 38.6

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 $X = \mu + \sigma Z$ then

$$\begin{aligned} M_X(t) &= E(e^{tX}) = E(e^{t\mu + t\sigma Z}) = E(e^{t\mu} e^{t\sigma Z}) = e^{t\mu} E(e^{t\sigma Z}) \\ &= e^{t\mu} M_Z(t\sigma) = e^{t\mu} e^{\frac{\sigma^2 t^2}{2}} = \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\} \end{aligned}$$

By differentiation we obtain

$$M'_X(t) = (\mu + t\sigma^2) \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\}$$

and

$$M''_X(t) = (\mu + t\sigma^2)^2 \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\} + \sigma^2 \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\}$$

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

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \sigma^2 \blacksquare$$

Next, suppose X_1, X_2, \dots, X_N are independent random variables. Then, the moment generating function of $Y = X_1 + \dots + X_N$ is

$$M_Y(t) = E(e^{t(X_1 + X_2 + \dots + X_N)}) = E(e^{X_1 t} \dots e^{X_N t})$$

$$= \prod_{k=1}^N E(e^{X_k t}) = \prod_{k=1}^N M_{X_k}(t)$$

where the next-to-last equality follows from Proposition 35.5.

Example 38.7

A company insures homes in three cities, J, K, L . 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 $E(X^3)$.

Solution.

Let J, K, L denote the losses from the three cities. Then $X = J + K + L$. Since J, K, L are independent, the moment-generating function for their sum, X , is equal to the product of the individual moment-generating functions, i.e.,

$$M_X(t) = M_K(t)M_J(t)M_L(t) = (1 - 2t)^{-3-2.5-4.5} = (1 - 2t)^{-10}$$

Differentiating this function, we get

$$\begin{aligned} M'(t) &= (-2)(-10)(1 - 2t)^{-11}; \\ M''(t) &= (-2)^2(-10)(-11)(1 - 2t)^{-12}; \\ M'''(t) &= (-2)^3(-10)(-11)(-12)(1 - 2t)^{-13} \end{aligned}$$

Hence, $E(X^3) = M_X'''(0) = (-2)^3(-10)(-11)(-12) = 10,560$. ■

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 $\{0, 1, 2, \dots, n\}$. 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 $i = 0, 1, \dots, n$. 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, \dots, c_n . The only way it can be zero for all values of s is if $c_0 = c_1 = \dots = c_n = 0$.

That is

$$p_X(i) = p_Y(i), \quad i = 0, 1, 2, \dots, n.$$

So probability mass functions for X and Y are exactly the same.

Example 38.8

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 38.9

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 38.10

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

$$\begin{aligned} M_{X+Y}(t) &= M_X(t)M_Y(t) \\ &= \exp\left\{\frac{\sigma_1^2 t^2}{2} + \mu_1 t\right\} \cdot \exp\left\{\frac{\sigma_2^2 t^2}{2} + \mu_2 t\right\} \\ &= \exp\left\{\frac{(\sigma_1^2 + \sigma_2^2)t^2}{2} + (\mu_1 + \mu_2)t\right\} \end{aligned}$$

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 38.11

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 $P(X_1 + X_2 + X_3 > X_4)$, which is the same as $P(X_1 + X_2 + X_3 - X_4 > 0)$. Now, since the X_i s 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

$$\begin{aligned} P(X > 0) &= P\left(\frac{X - 1}{\sqrt{9}} > \frac{0 - 1}{\sqrt{9}}\right) \\ &= P(Z > -0.33) = 1 - P(Z \leq -0.33) \\ &= P(Z \leq 0.33) \approx 0.6293 \blacksquare \end{aligned}$$

Joint Moment Generating Functions

For any random variables X_1, X_2, \dots, X_n , the joint moment generating function is defined by

$$M(t_1, t_2, \dots, t_n) = E(e^{t_1 X_1 + t_2 X_2 + \dots + t_n X_n}).$$

Example 38.12

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 38.13

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,

$$E(XY) = \frac{\partial^2}{\partial t_2 \partial t_1} M(t_1, t_2) \Big|_{(0,0)} = \frac{1}{(1-t_1)^2(1-t_2)^2} \Big|_{(0,0)} = 1$$

$$E(X) = \frac{\partial}{\partial t_1} M(t_1, t_2) \Big|_{(0,0)} = \frac{1}{(1-t_1)^2(1-t_2)} \Big|_{(0,0)} = 1$$

$$E(Y) = \frac{\partial}{\partial t_2} M(t_1, t_2) \Big|_{(0,0)} = \frac{1}{(1-t_1)(1-t_2)^2} \Big|_{(0,0)} = 1$$

and

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0 \blacksquare$$

Problems

Problem 38.1

Let X be a discrete random variable with range $\{1, 2, \dots, n\}$ 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 38.2

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 38.3

The following problem exhibits a random variable with no moment generating function. Let X be a random variable with pmf given by

$$p_X(n) = \frac{6}{\pi^2 n^2}, \quad n = 1, 2, 3, \dots$$

Show that $M_X(t)$ does not exist in any neighborhood of 0.

Problem 38.4

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 38.5

Show that the sum of n independently exponential random variable each with parameter λ is a gamma random variable with parameters n and λ .

Problem 38.6

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 38.7

Let X be an exponential random variable with parameter λ . Find the moment generating function of $Y = 3X - 2$.

Problem 38.8

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 38.9

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 38.10 ‡

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 38.11 ‡

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 38.12 ‡

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 38.13 ‡

Let X_1, X_2, X_3 be discrete random variables with common probability mass function

$$p(x) = \begin{cases} \frac{1}{3} & x = 0 \\ \frac{2}{3} & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Determine the moment generating function $M(t)$, of $Y = X_1X_2X_3$.

Problem 38.14 ‡

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 38.15

Let X_1, X_2, \dots, X_n be independent geometric random variables each with parameter p . Define $Y = X_1 + X_2 + \dots + X_n$.

- Find the moment generating function of X_i , $1 \leq i \leq n$.
- Find the moment generating function of a negative binomial random variable with parameters (n, p) .
- Show that Y defined above is a negative binomial random variable with parameters (n, p) .

Problem 38.16

Assume the math scores on the SAT test are normally distributed with mean 500 and standard deviation 60, and the verbal scores are normally distributed with mean 450 and standard deviation 80. If two students who took both tests are chosen at random, what is the probability that the first student's math score exceeds the second student's verbal score?

Problem 38.17

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 38.18

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 38.19

If the moment generating function for the random variable X is $M_X(t) = \frac{1}{t+1}$, find $E[(X-2)^3]$.

Problem 38.20

Suppose that X is a random variable with moment generating function $M_X(t) = \sum_{j=0}^{\infty} \frac{e^{(tj-1)}}{j!}$. Find $P(X=2)$.

Problem 38.21

If X has a standard normal distribution and $Y = e^X$, what is the k -th moment of Y ?

Problem 38.22

The random variable X has an exponential distribution with parameter b . It is found that $M_X(-b^2) = 0.2$. Find b .

Problem 38.23

Let X_1 and X_2 be two random variables with joint density function

$$f_{X_1X_2}(x_1, x_2) = \begin{cases} 1 & 0 < x_1 < 1, 0 < x_2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the moment generating function $M(t_1, t_2)$.

Problem 38.24

The moment generating function for the joint distribution of random variables X and Y is $M(t_1, t_2) = \frac{1}{3(1-t_2)} + \frac{2}{3}e^{t_1} \cdot \frac{2}{(2-t_2)}$, $t_2 < 1$. Find $\text{Var}(X)$.

Problem 38.25

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 38.26

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 38.27

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 38.28

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 38.29

The birth weight of males is normally distributed with mean 6 pounds, 10 ounces, standard deviation 1 pound. For females, the mean weight is 7 pounds, 2 ounces with standard deviation 12 ounces. Given two independent male/female births, find the probability that the baby boy outweighs the baby girl.

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.

39 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.

39.1 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 39.1 (*Markov's Inequality*)

If $X \geq 0$ and $c > 0$, then $P(X \geq c) \leq \frac{E(X)}{c}$.

Proof.

Let $c > 0$. Define

$$I = \begin{cases} 1 & \text{if } X \geq c \\ 0 & \text{otherwise} \end{cases}$$

Since $X \geq 0$ then $I \leq \frac{X}{c}$. Taking expectations of both side we find $E(I) \leq \frac{E(X)}{c}$. Now the result follows since $E(I) = P(X \geq c)$ ■

Example 39.1

From past experience a professor knows that the test score of a student taking her final examination 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 score on the final exam. Using Markov's inequality, we have

$$P(X \geq 85) \leq \frac{E(X)}{85} = \frac{75}{85} \approx 0.882 \blacksquare$$

Example 39.2

If X is a non-negative random variable, then for all $a > 0$ we have $P(X \geq aE(X)) \leq \frac{1}{a}$.

Solution.

The result follows by letting $c = aE(X)$ in Markov's inequality ■

Remark 39.1

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 $P(X = -1000) = P(X = 1000) = \frac{1}{2}$. Then $E(X) = 0$ and $P(X \geq 1000) \neq 0$

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 39.2

Suppose that X is a random variable such that $R \leq M$ for some constant M . Then for all $x < M$ we have

$$P(X \leq x) \leq \frac{M - E(X)}{M - x}$$

Proof.

By applying Markov's inequality we find

$$\begin{aligned} P(X \leq x) &= P(M - X \geq M - x) \\ &\leq \frac{E(M - X)}{M - x} = \frac{M - E(X)}{M - x} \blacksquare \end{aligned}$$

Example 39.3

Let X denote the test score of a random student, where the maximum score obtainable is 100. Find an upper bound of $P(X \leq 50)$, given that $E(X) = 75$.

Solution.

By the previous proposition we find

$$P(X \leq 50) \leq \frac{100 - 75}{100 - 50} = \frac{1}{2} \blacksquare$$

As a corollary of Proposition 39.1 we have

Proposition 39.3 (*Chebyshev's Inequality*)

If X is a random variable with finite mean μ and variance σ^2 , then for any value $\epsilon > 0$,

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

Proof.

Since $(X - \mu)^2 \geq 0$ then by Markov's inequality we can write

$$P[(X - \mu)^2 \geq \epsilon^2] \leq \frac{E[(X - \mu)^2]}{\epsilon^2}.$$

But $(X - \mu)^2 \geq \epsilon^2$ is equivalent to $|X - \mu| \geq \epsilon$ and this in turn is equivalent to

$$P(|X - \mu| \geq \epsilon) \leq \frac{E[(X - \mu)^2]}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2} \blacksquare$$

Example 39.4

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 39.5

Suppose X is the IQ of a random person. We assume $X \geq 0$, $E(X) = 100$, and $\sigma = 10$. Find an upper bound of $P(X \geq 200)$ using first Markov's inequality and then Chebyshev's inequality.

Solution.

By using Markov's inequality we find

$$P(X \geq 200) \leq \frac{100}{200} = \frac{1}{2}$$

Now, using Chebyshev's inequality we find

$$\begin{aligned} P(X \geq 200) &= P(X - 100 \geq 100) \\ &= P(X - E(X) \geq 10\sigma) \\ &\leq P(|X - E(X)| \geq 10\sigma) \leq \frac{1}{100} \blacksquare \end{aligned}$$

Example 39.6

On a single tank of gasoline, you expect to be able to drive 240 miles before running out of gas.

(a) Let p be the probability that you will NOT be able to make a 300 mile journey on a single tank of gas. What can you say about p ?

(b) Assume now that you also know that the standard deviation of the distance you can travel on a single tank of gas is 30 miles. What can you say now about p ?

Solution.

(a) Let X be the random variable representing mileage. Then by using Markov's inequality we find

$$p = P(X < 300) = 1 - P(X \geq 300) \geq 1 - \frac{240}{300} = 0.2$$

(b) By Chebyshev's inequality we find

$$p = P(X < 300) = 1 - P(X \geq 300) \geq 1 - P(|X - 240| \geq 60) \geq 1 - \frac{900}{3600} = 0.75 \blacksquare$$

Example 39.7

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 $P(|X - E(X)| \geq \frac{n}{3}) \leq \frac{9}{4n}$.

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

$$\begin{aligned} E(X^2) &= E \left[\left(\sum_{i=1}^n X_i \right)^2 \right] = nE(X_1^2) + \sum_{i \neq j} E(X_i X_j) \\ &= \frac{n}{2} + n(n-1) \frac{1}{4} = \frac{n(n+1)}{4} \end{aligned}$$

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:

$$P(|X - E(X)| \geq \frac{n}{3}) \leq \frac{\text{Var}(X)}{(n/3)^2} = \frac{9}{4n} \blacksquare$$

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 39.4

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 $P(X = 0) = 1$. Since $E(X) = 0$, by Markov's inequality $P(X \geq c) = 0$ for all $c > 0$. But

$$P(X > 0) = P \left(\bigcup_{n=1}^{\infty} (X > \frac{1}{n}) \right) \leq \sum_{n=1}^{\infty} P(X > \frac{1}{n}) = 0.$$

Hence, $P(X > 0) = 0$. Since $X \geq 0$, $1 = P(X \geq 0) = P(X = 0) + P(X > 0) = P(X = 0)$.

Now, suppose that $\text{Var}(X) = 0$. Since $(X - E(X))^2 \geq 0$ and $\text{Var}(X) = E((X - E(X))^2)$, by the above result we have $P(X - E(X) = 0) = 1$. That is, $P(X = E(X)) = 1 \blacksquare$

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 39.1

Let X_1, X_2, \dots, X_n be a sequence of independent random variables with common mean μ and finite common variance σ^2 . Then for any $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mu \right| \geq \epsilon \right\} = 0$$

or equivalently

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mu \right| < \epsilon \right) = 1$$

Proof.

Since

$$E \left[\frac{X_1 + X_2 + \dots + X_n}{n} \right] = \mu \text{ and } \text{Var} \left(\frac{X_1 + X_2 + \dots + X_n}{n} \right) = \frac{\sigma^2}{n}$$

by Chebyshev's inequality we find

$$0 \leq P \left\{ \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mu \right| \geq \epsilon \right\} \leq \frac{\sigma^2}{n\epsilon^2}$$

and the result follows by letting $n \rightarrow \infty$ ■

The above theorem says that for large n , $\frac{X_1 + X_2 + \dots + X_n}{n} - \mu$ is small with high probability. Also, it says that the distribution of the sample average becomes concentrated near μ as $n \rightarrow \infty$.

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 $S_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ 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

$$\lim_{n \rightarrow \infty} P(|S_n - \mu| < \epsilon) = 1$$

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 $p = P(A)$. This agrees with the definition of probability that we introduced in Section 6 p. 54.

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 39.8

Let X_1, X_2, \dots be pairwise independent random variables such that $\text{Var}(X_i) \leq b$ for some constant $b > 0$ and for all $1 \leq i \leq n$. Let

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

and

$$\mu_n = E(S_n).$$

Show that, for every $\epsilon > 0$ we have

$$P(|S_n - \mu_n| > \epsilon) \leq \frac{b}{\epsilon^2} \cdot \frac{1}{n}$$

and consequently

$$\lim_{n \rightarrow \infty} P(|S_n - \mu_n| \leq \epsilon) = 1$$

Solution.

Since $E(S_n) = \mu_n$ and $\text{Var}(S_n) = \frac{\text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n)}{n^2} \leq \frac{b}{n^2}$, by Chebyshev's inequality we find

$$0 \leq P\left\{\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mu_n\right| \geq \epsilon\right\} \leq \frac{b}{n^2} \cdot \frac{1}{n}.$$

Now,

$$1 \geq P(|S_n - \mu_n| \leq \epsilon) = 1 - P(|S_n - \mu_n| > \epsilon) \geq 1 - \frac{b}{\epsilon^2} \cdot \frac{1}{n}$$

By letting $n \rightarrow \infty$ we conclude that

$$\lim_{n \rightarrow \infty} P(|S_n - \mu_n| \leq \epsilon) = 1 \blacksquare$$

39.2 The Strong Law of Large Numbers

Recall the weak law of large numbers:

$$\lim_{n \rightarrow \infty} P(|S_n - \mu| < \epsilon) = 1$$

where the X_i 's are independent identically distributed random variables and $S_n = \frac{X_1 + X_2 + \dots + X_n}{n}$. 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 $\lim_{n \rightarrow \infty} S_n(x) = \mu$. Fortunately, however, there is a stronger version of the law of large numbers that does assure convergence for individual realizations.

Theorem 39.2

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

$$P\left(\lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right) = 1.$$

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

$$X_i^4, \quad X_i^3 X_j, \quad X_i^2 X_j^2, \quad X_i^2 X_j X_k \quad \text{and} \quad X_i X_j X_k X_l$$

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

$$\begin{aligned} E(X_i^3 X_j) &= E(X_i^3)E(X_j) = 0 \\ E(X_i^2 X_j X_k) &= E(X_i^2)E(X_j)E(X_k) = 0 \\ E(X_i X_j X_k X_l) &= E(X_i)E(X_j)E(X_k)E(X_l) = 0 \end{aligned}$$

Next, there are n terms of the form X_i^4 and for each $i \neq j$ the coefficient of $X_i^2 X_j^2$ according to the multinomial theorem is

$$\frac{4!}{2!2!} = 6.$$

But there are $C(n, 2) = \frac{n(n-1)}{2}$ 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

$$0 \leq \text{Var}(X_i^2) = E(X_i^4) - (E(X_i^2))^2$$

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.

$$E\left[\sum_{n=1}^{\infty} \frac{T_n^4}{n^4}\right] \leq 4K \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty \quad (39.1)$$

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 (39.1). 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 n th term goes to 0. Hence,

$$P\left[\lim_{n \rightarrow \infty} \frac{T_n^4}{n^4} = 0\right] = 1$$

Since $\frac{T_n^4}{n^4} = (T_n/n)^4 = S_n^4$ the last result implies that

$$P\left[\lim_{n \rightarrow \infty} S_n = 0\right] = 1$$

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

$$P\left[\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{(X_i - \mu)}{n} = 0\right] = 1$$

or equivalently

$$P \left[\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{X_i}{n} = \mu \right] = 1$$

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 $P(E)$ denote the probability that E occurs on any particular trial. Define

$$X_i = \begin{cases} 1 & \text{if } E \text{ occurs on the } i\text{th trial} \\ 0 & \text{otherwise} \end{cases}$$

By the Strong Law of Large Numbers we have

$$P \left[\lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = E(X) = P(E) \right] = 1$$

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 $P(E)$. This justifies our definition of $P(E)$ that we introduced in Section 6, i.e.,

$$P(E) = \lim_{n \rightarrow \infty} \frac{n(E)}{n}$$

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 39.9

Let X_1, X_2, \cdots be the sequence of mutually independent random variables such that $X_1 = 0$, and for each positive integer i

$$P(X_i = i) = \frac{1}{2i \ln i}, \quad P(X_i = -i) = \frac{1}{2i \ln i}, \quad P(X_i = 0) = 1 - \frac{1}{i \ln i}$$

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 $\text{Var}(T_n) \leq \frac{n^2}{\ln n}$.

(b) Show that the sequence X_1, X_2, \dots satisfies the Weak Law of Large Numbers, i.e., prove that for any $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P(|S_n| \geq \epsilon) = 0.$$

(c) Let A_1, A_2, \dots be any infinite sequence of mutually independent events such that

$$\sum_{i=1}^{\infty} P(A_i) = \infty.$$

Prove that

$$P(\text{infinitely many } A_i \text{ occur}) = 1$$

(d) Show that $\sum_{i=1}^{\infty} P(|X_i| \geq i) = \infty$.

(e) Conclude that

$$\lim_{n \rightarrow \infty} P(S_n = \mu) = 0$$

and hence that the Strong Law of Large Numbers completely fails for the sequence X_1, X_2, \dots

Solution.

(a) We have

$$\begin{aligned} \text{Var}(T_n) &= \text{Var}(X_1) + \text{Var}(X_2) + \cdots + \text{Var}(X_n) \\ &= 0 + \sum_{i=2}^n (E(X_i^2) - (E(X_i))^2) \\ &= \sum_{i=2}^n i^2 \frac{1}{i \ln i} \\ &= \sum_{i=2}^n \frac{i}{\ln i} \end{aligned}$$

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 $2 \leq i \leq n$. Furthermore,

$$\frac{n^2}{\ln n} = \sum_{i=1}^n \frac{n}{\ln n} \geq \sum_{i=2}^n \frac{i}{\ln i}$$

Hence,

$$\text{Var}(T_n) \leq \frac{n^2}{\ln n}.$$

(b) We have

$$\begin{aligned} P(|S_n| \geq \epsilon) &= P(|S_n - 0| \geq \epsilon) \\ &\leq \text{Var}(S_n) \cdot \frac{1}{\epsilon^2} \quad \text{Chebyshev's inequality} \\ &= \frac{\text{Var}(T_n)}{n^2} \cdot \frac{1}{\epsilon^2} \\ &\leq \frac{1}{\epsilon^2 \ln n} \end{aligned}$$

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

$$\lim_{n \rightarrow \infty} E(T_{r,n}) = \lim_{n \rightarrow \infty} \sum_{i=r}^n E(I_{A_i}) = \lim_{n \rightarrow \infty} \sum_{i=r}^n P(A_i) = \infty$$

Since $e^x \rightarrow 0$ as $x \rightarrow -\infty$ we conclude that $e^{-E(T_{r,n})} \rightarrow 0$ as $n \rightarrow \infty$.

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 occur. We must prove that $P(K) = 0$. We first show that

$$P(K_{r,n}) \leq e^{-E(T_{r,n})}.$$

We remind the reader of the inequality $1 + x \leq e^x$. We have

$$\begin{aligned}
 P(K_{r,n}) &= P(T_{r,n} = 0) = P[(A_r \cup A_{r+1} \cup \cdots \cup A_n)^c] \\
 &= P[A_r^c \cap A_{r+1}^c \cap \cdots \cap A_n^c] \\
 &= \prod_{i=r}^n P(A_i^c) \\
 &= \prod_{i=r}^n [1 - P(A_i)] \\
 &\leq \prod_{i=r}^n e^{-P(A_i)} \\
 &= e^{-\sum_{i=r}^n P(A_i)} \\
 &= e^{-\sum_{i=r}^n E(I_{A_i})} \\
 &= e^{-E(T_{r,n})}
 \end{aligned}$$

Now, since $K_r \subset K_{r,n}$ we conclude that $0 \leq P(K_r) \leq P(K_{r,n}) \leq e^{-E(T_{r,n})} \rightarrow 0$ as $n \rightarrow \infty$. Hence, $P(K_r) = 0$ for all $r \leq n$.

Now note that $K = \cup_r K_r$ so by Boole's inequality (See Proposition 35.3), $0 \leq P(K) \leq \sum_r P(K_r) = 0$. That is, $P(K) = 0$. Hence the probability that infinitely many A_i 's occurs is 1.

(d) We have that $P(|X_i| \geq i) = \frac{1}{i \ln i}$. Thus,

$$\begin{aligned}
 \sum_{i=1}^n P(|X_i| \geq i) &= 0 + \sum_{i=2}^n \frac{1}{i \ln i} \\
 &\geq \int_2^n \frac{dx}{x \ln x} \\
 &= \ln \ln n - \ln \ln 2
 \end{aligned}$$

and this last term approaches infinity and 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 $\lim_{n \rightarrow \infty} \frac{X_n}{n} \neq 0$. Hence,

$$P\left(\lim_{n \rightarrow \infty} \frac{X_n}{n} \neq 0\right) = 1$$

which means

$$P\left(\lim_{n \rightarrow \infty} \frac{X_n}{n} = 0\right) = 0$$

Now note that

$$\frac{X_n}{n} = S_n - \frac{n-1}{n}S_{n-1}$$

so that if $\lim_{n \rightarrow \infty} S_n = 0$ then $\lim_{n \rightarrow \infty} \frac{X_n}{n} = 0$. This implies that

$$P(\lim_{n \rightarrow \infty} S_n = 0) \leq P(\lim_{n \rightarrow \infty} \frac{X_n}{n} = 0)$$

That is,

$$P(\lim_{n \rightarrow \infty} S_n = 0) = 0$$

and this violates the Strong Law of Large numbers ■

Problems

Problem 39.1

Let $\epsilon > 0$. Let X be a discrete random variable with range $\{-\epsilon, \epsilon\}$ and pmf given by $p(-\epsilon) = \frac{1}{2}$ and $p(\epsilon) = \frac{1}{2}$ and 0 otherwise. Show that the inequality in Chebyshev's inequality is in fact an equality.

Problem 39.2

Let X_1, X_2, \dots, X_n be a Bernoulli trials process with probability .3 for success and .7 for failure. Let $X_i = 1$ if the i th outcome is a success and 0 otherwise. Find n so that $P(|\frac{S_n}{n} - E(\frac{S_n}{n})| \geq 0.1) \leq 0.21$, where $S_n = X_1 + X_2 + \dots + X_n$.

Problem 39.3

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 39.4

Consider the transmission of several 10Mbytes files over a noisy channel. Suppose that the average number of erroneous bits per transmitted file at the receiver output is 10^3 . What can be said about the probability of having $\geq 10^4$ erroneous bits during the transmission of one of these files?

Problem 39.5

Suppose that X is a random variable with mean and variance both equal to 20. What can be said about $P(0 < X < 40)$?

Problem 39.6

Let X_1, X_2, \dots, X_{20} be independent Poisson random variables with mean 1. Use Markov's inequality to obtain a bound on

$$P(X_1 + X_2 + \dots + X_{20} > 15).$$

Problem 39.7

From past experience a professor knows that the test score of a student taking her final examination is a random variable with mean 75 and variance 25. What can be said about the probability that a student will score between 65 and 85?

Problem 39.8

Let $M_X(t) = E(e^{tX})$ be the moment generating function of a random variable X . Show that

$$P(X \geq \epsilon) \leq e^{-\epsilon t} M_X(t).$$

Problem 39.9

On average a shop sells 50 telephones each week. Stock one Monday is 75. What can be said about the probability of not having enough stock left by the end of the week?

Problem 39.10

Let X be a random variable with mean $\frac{1}{2}$ and variance 25×10^{-7} . What can you say about $P(0.475 \leq X \leq 0.525)$.

Problem 39.11

Let X be a random variable with mean μ . Show that $P(X \geq 2\mu) \leq \frac{1}{2}$.

Problem 39.12

The numbers of robots produced in a factory during a day is a random variable with mean 100. Describe the probability that the factory's production will be more than 120 in a day. Also if the variance is known to be 5, then describe the probability that the factory's production will be between 70 and 130 in a day.

Problem 39.13

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.
- Use normal approximation to compute the probability that X is between 100 and 140.

Problem 39.14

Let X_1, \dots, X_n be independent random variables, each with probability density function:

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

Show that $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ converges in probability to a constant as $n \rightarrow \infty$ and find that constant.

Problem 39.15

Let X_1, \dots, X_n be independent and identically distributed Uniform(0,1). Let Y_n be the minimum of X_1, \dots, 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 $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P(|Y_n - 0| \leq \epsilon) = 1.$$

40 The Central Limit Theorem

The central limit theorem is one of the most remarkable theorem among the limit theorems. This theorem says that the sum of large number of independent identically distributed random variables is well-approximated by a normal random variable.

We first need a technical result.

Theorem 40.1

Let Z_1, Z_2, \dots 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 $M_{Z_n}(t) \rightarrow M_Z(t)$ as $n \rightarrow \infty$ and for all t , then $F_{Z_n} \rightarrow F_Z$ for all t at which $F_Z(t)$ is continuous.

With the above theorem, we can prove the central limit theorem.

Theorem 40.2

Let X_1, X_2, \dots be a sequence of independent and identically distributed random variables, each with mean μ and variance σ^2 . Then,

$$P\left(\frac{\sqrt{n}}{\sigma}\left(\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right) \leq a\right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{x^2}{2}} dx$$

as $n \rightarrow \infty$.

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 $\frac{\sqrt{n}}{\sigma}\left(\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right)$ 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

$$M_{\frac{\sqrt{n}}{\sigma}\left(\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right)}(x) \rightarrow e^{\frac{x^2}{2}}$$

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

$$\begin{aligned}
 M_{\frac{\sqrt{n}}{\sigma} \left(\frac{X_1 + X_2 + \dots + X_n}{n} - \mu \right)}(x) &= M_{\frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - \mu}{\sigma}}(x) \\
 &= M_{\sum_{i=1}^n \frac{X_i - \mu}{\sigma}} \left(\frac{x}{\sqrt{n}} \right) \\
 &= \prod_{i=1}^n M_{\frac{X_i - \mu}{\sigma}} \left(\frac{x}{\sqrt{n}} \right) \\
 &= \left(M_{\frac{X_1 - \mu}{\sigma}} \left(\frac{x}{\sqrt{n}} \right) \right)^n \\
 &= \left(M_Y \left(\frac{x}{\sqrt{n}} \right) \right)^n
 \end{aligned}$$

where

$$Y = \frac{X_1 - \mu}{\sigma}.$$

Now expand $M_Y(x/\sqrt{n})$ in a Taylor series around 0 as follows

$$M_Y \left(\frac{x}{\sqrt{n}} \right) = M_Y(0) + M'_Y(0) \left(\frac{x}{\sqrt{n}} \right) + \frac{1}{2} \frac{x^2}{n} M''_Y(0) + R \left(\frac{x}{\sqrt{n}} \right), \quad |x| < \sqrt{n}\sigma h$$

where

$$\frac{n}{x^2} R \left(\frac{x}{\sqrt{n}} \right) \rightarrow 0 \text{ as } x \rightarrow 0$$

But

$$E(Y) = E \left[\frac{X_1 - \mu}{\sigma} \right] = 0$$

and

$$E(Y^2) = E \left[\left(\frac{X_1 - \mu}{\sigma} \right)^2 \right] = \frac{\text{Var}(X_1)}{\sigma^2} = 1.$$

By Proposition 38.1 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

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(M_Y \left(\frac{x}{\sqrt{n}} \right) \right)^n &= \lim_{n \rightarrow \infty} \left(1 + \frac{1}{2} \frac{x^2}{n} + R \left(\frac{x}{\sqrt{n}} \right) \right)^n \\ &= \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \left(\frac{x^2}{2} + nR \left(\frac{x}{\sqrt{n}} \right) \right) \right)^n \end{aligned}$$

But

$$nR \left(\frac{x}{\sqrt{n}} \right) = x^2 \frac{n}{x^2} R \left(\frac{x}{\sqrt{n}} \right) \rightarrow 0$$

as $n \rightarrow \infty$. Hence,

$$\lim_{n \rightarrow \infty} \left(M_Y \left(\frac{x}{\sqrt{n}} \right) \right)^n = e^{\frac{x^2}{2}}.$$

Now the result follows from Theorem 40.1 with $Z_n = \frac{\sqrt{n}}{\sigma} \left(\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right)$, Z standard normal distribution,

$$F_{Z_n}(a) = P \left(\frac{\sqrt{n}}{\sigma} \left(\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right) \leq a \right)$$

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 40.1

The weight of an arbitrary airline passenger's baggage has a mean of 20 pounds and a variance of 9 pounds. Consider an airplane that carries 200 passengers, and assume every passenger checks one piece of luggage. Estimate the probability that the total baggage weight exceeds 4050 pounds.

Solution.

Let X_i = weight of baggage of passenger i . Thus,

$$\begin{aligned} P\left(\sum_{i=1}^{200} X_i > 4050\right) &= P\left(\frac{\sum_{i=1}^{200} X_i - 200(20)}{3\sqrt{200}} > \frac{4050 - 20(200)}{3\sqrt{200}}\right) \\ &\approx P(Z > 1.179) = 1 - P(Z \leq 1.179) \\ &= 1 - \Phi(1.179) = 0.1314 \blacksquare \end{aligned}$$

Example 40.2 †

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 \bar{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,

$$\begin{aligned} P\left(-\frac{1}{4} \leq \bar{X}_{48} \leq \frac{1}{4}\right) &= P\left(\frac{-0.25}{0.2083} \leq \frac{X_{48}}{0.2083} \leq \frac{0.25}{0.2083}\right) \\ &= P(-1.2 \leq Z \leq 1.2) = 2\Phi(1.2) - 1 \approx 2(0.8849) - 1 = 0.77 \blacksquare \end{aligned}$$

Example 40.3

A computer generates 48 random real numbers, rounds each number to the nearest integer and then computes the average of these 48 rounded values. Assume that the numbers generated are independent of each other and that the rounding errors are distributed uniformly on the interval $[-0.5, 0.5]$. Find the approximate probability that the average of the rounded values is within 0.05 of the average of the exact numbers.

Solution.

Let X_1, \dots, X_{48} denote the 48 rounding errors, and $\bar{X} = \frac{1}{48} \sum_{i=1}^{48} X_i$ their average. We need to compute $P(|\bar{X}| \leq 0.05)$. Since a rounding error 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, \bar{X} has approximate distribution $N(\mu, \frac{\sigma^2}{n}) = N(0, \frac{1}{24^2})$. Thus $24\bar{X}$ is approximately standard normal, so

$$\begin{aligned} P(|\bar{X}| \leq 0.05) &\approx P(24 \cdot (-0.05) \leq 24\bar{X} \leq 24 \cdot (0.05)) \\ &= \Phi(1.2) - \Phi(-1.2) = 2\Phi(1.2) - 1 = 0.7698 \blacksquare \end{aligned}$$

Example 40.4

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 \bar{X} be the sample mean. Determine a such that $P(\bar{X} \leq a) = 0.90$.

Solution.

The sample mean \bar{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

$$0.90 = P(\bar{X} \leq a) = P\left(\frac{\bar{X} - 2}{1.58} < \frac{a - 2}{1.58}\right) = \Phi\left(\frac{a - 2}{1.58}\right).$$

From the normal table, we get $\frac{a-2}{1.58} = 1.28$, so $a = 4.02$ ■

Example 40.5

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.

- What is the probability that the load (total weight) exceeds the design limit?
- 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 $Y = X_1 + X_2 + \dots + X_{25}$, 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

$$\begin{aligned} P(Y > 4300) &= P\left(\frac{Y - 4000}{\sqrt{22500}} > \frac{4300 - 4000}{\sqrt{22500}}\right) \\ &= P(Z > 2) = 1 - P(Z \leq 2) \\ &= 1 - 0.9772 = 0.0228 \end{aligned}$$

(b) We want to find x such that $P(Y > x) = 0.001$. Note that

$$\begin{aligned} P(Y > x) &= P\left(\frac{Y - 4000}{\sqrt{22500}} > \frac{x - 4000}{\sqrt{22500}}\right) \\ &= P\left(Z > \frac{x - 4000}{\sqrt{22500}}\right) = 0.01 \end{aligned}$$

It is equivalent to $P\left(Z \leq \frac{x-4000}{\sqrt{22500}}\right) = 0.999$. From the normal Table we find $P(Z \leq 3.09) = 0.999$. So $(x - 4000)/150 = 3.09$. Solving for x we find $x \approx 4463.5$ pounds ■

Problems**Problem 40.1**

A manufacturer of booklets packages the booklets in boxes of 100. It is known that, on average, the booklets weigh 1 ounce, with a standard deviation of 0.05 ounces. What is the probability that 1 box of booklets weighs more than 100.4 ounces?

Problem 40.2

In a particular basketball league, the standard deviation in the distribution of players' height is 2 inches. Twenty-five players are selected at random and their heights are measured. Give an estimate for the probability that the average height of the players in this sample of 25 is within 1 inch of the league average height.

Problem 40.3

A lightbulb manufacturer claims that the lifespan of its lightbulbs has a mean of 54 months and a st. deviation of 6 months. Your consumer advocacy group tests 50 of them. Assuming the manufacturer's claims are true, what is the probability that it finds a mean lifetime of less than 52 months?

Problem 40.4

If 10 fair dice are rolled, find the approximate probability that the sum obtained is between 30 and 40, inclusive.

Problem 40.5

Let $X_i, i = 1, 2, \dots, 10$ be independent random variables each uniformly distributed over $(0,1)$. Calculate an approximation to $P(\sum_{i=1}^{10} X_i > 6)$.

Problem 40.6

Suppose that $X_i, i = 1, \dots, 100$ are exponentially distributed random variables with parameter $\lambda = \frac{1}{10000}$. Let $\bar{X} = \frac{\sum_{i=1}^{100} X_i}{100}$. Approximate $P(950 \leq \bar{X} \leq 1050)$.

Problem 40.7

The Braves play 100 independent baseball games, each of which they have probability 0.8 of winning. What's the probability that they win at least 90?

Problem 40.8

An insurance company has 10,000 automobile policyholders. The expected yearly claim per policyholder is \$240 with a standard deviation of \$800. Approximate the probability that the total yearly claim exceeds \$2.7 million.

Problem 40.9

A certain component is critical to the operation of an electrical system and must be replaced immediately upon failure. If the mean life of this type of component is 100 hours and its standard deviation is 30 hours, how many of these components must be in stock so that the probability that the system is in continual operation for the next 2000 hours is at least 0.95?

Problem 40.10

Student scores on exams given by a certain instructor has mean 74 and standard deviation 14. This instructor is about to give two exams. One to a class of size 25 and the other to a class of 64. Approximate the probability that the average test score in the class of size 25 exceeds 80.

Problem 40.11 ‡

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 40.12 ‡

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 40.13 ‡

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 40.14 ‡

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 :

$$\begin{aligned} E(X) &= 50 \\ E(Y) &= 20 \\ \text{Var}(X) &= 50 \\ \text{Var}(Y) &= 30 \\ \text{Cov}(X, Y) &= 10 \end{aligned}$$

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 $P(T < 7100)$.

Problem 40.15 ‡

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 40.16 ‡

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 40.17

Delivery times for shipments from a central warehouse are exponentially distributed with a mean of 2.4 days (note that times are measured continuously, not just in number of days). A random sample of $n = 100$ shipments are selected, and their shipping times are observed. What is the approximate probability that the average shipping time is less than 2.0 days?

Problem 40.18

(a) Give the approximate sampling distribution for the following quantity based on random samples of independent observations:

$$\bar{X} = \frac{\sum_{i=1}^{100} X_i}{100}, \quad E(X_i) = 100, \quad \text{Var}(X_i) = 400.$$

(b) What is the approximate probability the sample mean will be between 96 and 104?

41 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 41.1

Let X be a random variable with mean μ and finite variance σ^2 . Then for any $a > 0$

$$P(X \geq \mu + a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

and

$$P(X \leq \mu - a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

Proof.

Without loss of generality we assume that $\mu = 0$. Then for any $b > 0$ we have

$$\begin{aligned} P(X \geq a) &= P(X + b \geq a + b) \\ &\leq P((X + b)^2 \geq (a + b)^2) \\ &\leq \frac{E[(X + b)^2]}{(a + b)^2} \\ &= \frac{\sigma^2 + b^2}{(a + b)^2} \\ &= \frac{\alpha + t^2}{(1 + t)^2} = g(t) \end{aligned}$$

where

$$\alpha = \frac{\sigma^2}{a^2} \text{ and } t = \frac{b}{a}.$$

Since

$$g'(t) = \frac{1}{2} \frac{t^2 + (1 - \alpha)t - \alpha}{(1 + t)^4}$$

then $g'(t) = 0$ when $t = \alpha$. Since $g''(t) = 2(2t + 1 - \alpha)(1 + t)^{-4} - 8(t^2 + (1 - \alpha)t - \alpha)(1 + t)^{-5}$ then $g''(\alpha) = 2(\alpha + 1)^{-3} > 0$ 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

$$P(X \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}.$$

Now, if $\mu \neq 0$ then since $E(X - E(X)) = 0$ and $Var(X - E(X)) = Var(X) = \sigma^2$ then by applying the previous inequality to $X - \mu$ we obtain

$$P(X \geq \mu + a) \leq \frac{\sigma^2}{\sigma^2 + a^2}.$$

Similarly, since $E(\mu - X) = 0$ and $Var(\mu - X) = Var(X) = \sigma^2$ then we get

$$P(\mu - X \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

or

$$P(X \leq \mu - a) \leq \frac{\sigma^2}{\sigma^2 + a^2} \blacksquare$$

Example 41.1

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

$$P(X \geq 120) = P(X - 100 \geq 20) \leq \frac{400}{400 + 20^2} = \frac{1}{2} \blacksquare$$

The following provides bounds on $P(X \geq a)$ in terms of the moment generating function $M(t) = E(e^{tX})$ with $t > 0$.

Proposition 41.2 (*Chernoff's bound*)

Let X be a random variable and suppose that $M(t) = E(e^{tX})$ is finite. Then

$$P(X \geq a) \leq e^{-ta} M(t), \quad t > 0$$

and

$$P(X \leq a) \leq e^{-ta} M(t), \quad t < 0$$

Proof.

Suppose first that $t > 0$. Then

$$\begin{aligned} P(X \geq a) &\leq P(e^{tX} \geq e^{ta}) \\ &\leq E[e^{tX}]e^{-ta} \end{aligned}$$

where the last inequality follows from Markov's inequality. Similarly, for $t < 0$ we have

$$\begin{aligned} P(X \leq a) &\leq P(e^{tX} \geq e^{ta}) \\ &\leq E[e^{tX}]e^{-ta} \blacksquare \end{aligned}$$

It follows from Chernoff's inequality that a sharp bound for $P(X \geq a)$ is a minimizer of the function $e^{-ta}M(t)$.

Example 41.2

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 $P(Z \geq a)$.

Solution.

By Chernoff inequality we have

$$P(Z \geq a) \leq e^{-ta}e^{\frac{t^2}{2}} = e^{\frac{t^2}{2}-ta}, \quad t > 0$$

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 $g''(t) = e^{\frac{t^2}{2}-ta} + (t-a)^2e^{\frac{t^2}{2}-ta}$ then $g''(a) > 0$ so that $t = a$ is the minimum of $g(t)$. Hence, a sharp bound is

$$P(Z \geq a) \leq e^{-\frac{a^2}{2}}, \quad t > 0$$

Similarly, for $a < 0$ we find

$$P(Z \leq a) \leq e^{-\frac{a^2}{2}}, \quad t < 0 \blacksquare$$

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 41.3 (*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 41.3

Let X be a random variable. Show that $E(e^X) \geq e^{E(X)}$.

Solution.

Since $f(x) = e^x$ is convex then by Jensen's inequality we can write $E(e^X) \geq e^{E(X)}$ ■

Example 41.4

Suppose that $\{x_1, x_2, \dots, x_n\}$ 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 $p(X = x_i) = \frac{1}{n}$ 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 ■

Problems

Problem 41.1

Roll a single fair die and let X be the outcome. Then, $E(X) = 3.5$ and $Var(X) = \frac{35}{12}$.

- Compute the exact value of $P(X \geq 6)$.
- Use Markov's inequality to find an upper bound of $P(X \geq 6)$.
- Use Chebyshev's inequality to find an upper bound of $P(X \geq 6)$.
- Use one-sided Chebyshev's inequality to find an upper bound of $P(X \geq 6)$.

Problem 41.2

Find Chernoff bounds for a binomial random variable with parameters (n, p) .

Problem 41.3

Suppose that the average number of parking tickets given in the Engineering lot at Stony Brook is three per day. Also, assume that you are told that the variance of the number of tickets in any one day is 9. Give an estimate of the probability p , that at least five parking tickets will be given out in the Engineering lot tomorrow.

Problem 41.4

Each time I write a check, I omit the cents when I record the amount in my checkbook (I record only the integer number of dollars). Assume that I write 20 checks this month. Find an upper bound on the probability that the record in my checkbook shows at least \$15 less than the actual amount in my account.

Problem 41.5

Find the chernoff bounds for a Poisson random variable X with parameter λ .

Problem 41.6

Let X be a Poisson random variable with mean 20.

- Use the Markov's inequality to obtain an upper bound on

$$p = P(X \geq 26).$$

- Use the Chernoff bound to obtain an upper bound on p .
- Use the Chebyshev's bound to obtain an upper bound on p .
- Approximate p by making use of the central limit theorem.

Problem 41.7

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 41.8

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 41.9

Suppose that $\{x_1, x_2, \dots, x_n\}$ is a set of positive numbers. Prove

$$(x_1 \cdot x_2 \cdots x_n)^{\frac{2}{n}} \leq \frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}.$$

Appendix

42 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 $x = 0$.

Recall that the definite integral $\int_a^b f(x) dx$ was defined as the limit of a left- or right Riemann sum. We noted that the definite integral is always well-defined if:

- (a) $f(x)$ is continuous on $[a, b]$,
- (b) and if the domain of integration $[a, b]$ is finite.

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 :

$$\lim_{x \rightarrow c} f(x) = \pm\infty.$$

e.g.:

$$\int_0^1 \frac{1}{x} dx.$$

An improper integral may not be well defined or may have infinite value. In this case we say that the integral is **divergent**. In case an improper integral has a finite value then we say that it is **convergent**.

We will consider only improper integrals with positive integrands since they are the most common.

• Unbounded Intervals of Integration

The first type of improper integrals arises when the domain of integration is infinite. In case one of the limits of integration is infinite, we define

$$\int_a^\infty f(x) dx = \lim_{b \rightarrow \infty} \int_a^b f(x) dx$$

or

$$\int_{-\infty}^b f(x) dx = \lim_{a \rightarrow -\infty} \int_a^b f(x) dx.$$

If both limits are infinite, then we choose any number c in the domain of f and define

$$\int_{-\infty}^\infty f(x) dx = \int_{-\infty}^c f(x) dx + \int_c^\infty f(x) dx.$$

In this case, the integral is convergent if and only if both integrals on the right converge. Alternatively, we can also write

$$\int_{-\infty}^\infty f(x) dx = \lim_{R \rightarrow \infty} \int_{-R}^R f(x) dx.$$

Example 42.1

Does the integral $\int_1^\infty \frac{1}{x^2} dx$ converge or diverge?

Solution.

We have

$$\int_1^\infty \frac{1}{x^2} dx = \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^2} dx = \lim_{b \rightarrow \infty} \left[-\frac{1}{x} \right]_1^b = \lim_{b \rightarrow \infty} \left(-\frac{1}{b} + 1 \right) = 1.$$

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 42.1.

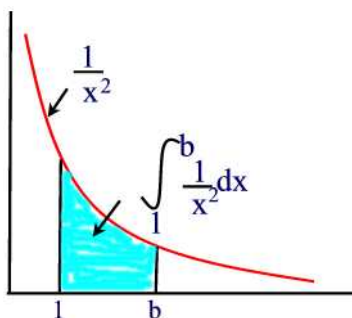


Figure 42.1

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 42.2

Does the improper integral $\int_1^\infty \frac{1}{\sqrt{x}} dx$ converge or diverge?

Solution.

We have

$$\int_1^\infty \frac{1}{\sqrt{x}} dx = \lim_{b \rightarrow \infty} \int_1^b \frac{1}{\sqrt{x}} dx = \lim_{b \rightarrow \infty} [2\sqrt{x}]_1^b = \lim_{b \rightarrow \infty} (2\sqrt{b} - 2) = \infty.$$

So the improper integral is divergent. ■

Remark 42.1

In general, some unbounded regions have finite areas while others have infinite areas. This is true whether a region extends to infinity along the x -axis or along the y -axis or both, and whether it extends to infinity on one or both sides of an axis. For example the area under the graph of $\frac{1}{x^2}$ from 1 to infinity is finite whereas that under the graph of $\frac{1}{\sqrt{x}}$ is infinite. This has to do with how fast each function approaches 0 as $x \rightarrow \infty$. The function $f(x) = x^2$ grows very rapidly which means that the graph is steep. When we consider the reciprocal $\frac{1}{x^2}$, it thus approaches the x axis very quickly and so the area bounded by this graph is finite.

On the other hand, the function $f(x) = x^{\frac{1}{2}}$ grows very slowly meaning that its graph is relatively flat. This means that the graph $y = \frac{1}{x^{\frac{1}{2}}}$ approaches the x axis very slowly and the area bounded by the graph is infinite.

The following example generalizes the results of the previous two examples.

Example 42.3

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

$$\begin{aligned} \int_1^\infty \frac{1}{x} dx &= \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x} dx \\ &= \lim_{b \rightarrow \infty} [\ln |x|]_1^b = \lim_{b \rightarrow \infty} \ln b = \infty \end{aligned}$$

so the improper integral is divergent.

Now, suppose that $p \neq 1$. Then

$$\begin{aligned} \int_1^\infty \frac{1}{x^p} dx &= \lim_{b \rightarrow \infty} \int_1^b x^{-p} dx \\ &= \lim_{b \rightarrow \infty} \left[\frac{x^{-p+1}}{-p+1} \right]_1^b \\ &= \lim_{b \rightarrow \infty} \left(\frac{b^{-p+1}}{-p+1} - \frac{1}{-p+1} \right). \end{aligned}$$

If $p < 1$ then $-p + 1 > 0$ so that $\lim_{b \rightarrow \infty} b^{-p+1} = \infty$ and therefore the improper integral is divergent. If $p > 1$ then $-p + 1 < 0$ so that $\lim_{b \rightarrow \infty} b^{-p+1} = 0$ and hence the improper integral converges:

$$\int_1^\infty \frac{1}{x^p} dx = \frac{-1}{-p+1}. \blacksquare$$

Example 42.4

For what values of c is the improper integral $\int_0^\infty e^{cx} dx$ convergent?

Solution.

We have

$$\begin{aligned} \int_0^\infty e^{cx} dx &= \lim_{b \rightarrow \infty} \int_0^b e^{cx} dx = \lim_{b \rightarrow \infty} \frac{1}{c} e^{cx} \Big|_0^b \\ &= \lim_{b \rightarrow \infty} \frac{1}{c} (e^{cb} - 1) = -1 \end{aligned}$$

provided that $c < 0$. Otherwise, i.e. if $c \geq 0$, then the improper integral is divergent. \blacksquare

Remark 42.2

The previous two results are very useful and you may want to memorize them.

Example 42.5

Show that the improper integral $\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx$ converges.

Solution.

Splitting the integral into two as follows:

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \int_{-\infty}^0 \frac{1}{1+x^2} dx + \int_0^{\infty} \frac{1}{1+x^2} dx.$$

Now,

$$\begin{aligned} \int_{-\infty}^0 \frac{1}{1+x^2} dx &= \lim_{a \rightarrow -\infty} \int_a^0 \frac{1}{1+x^2} dx = \lim_{a \rightarrow -\infty} \arctan x \Big|_a^0 \\ &= \lim_{a \rightarrow -\infty} (\arctan 0 - \arctan a) = -(-\frac{\pi}{2}) = \frac{\pi}{2}. \end{aligned}$$

Similarly, we find that $\int_0^{\infty} \frac{1}{1+x^2} dx = \frac{\pi}{2}$ so that $\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \frac{\pi}{2} + \frac{\pi}{2} = \pi$. ■

• Unbounded Integrands

Suppose $f(x)$ is unbounded at $x = a$, that is $\lim_{x \rightarrow a^+} f(x) = \infty$. Then we define

$$\int_a^b f(x) dx = \lim_{t \rightarrow a^+} \int_t^b f(x) dx.$$

Similarly, if $f(x)$ is unbounded at $x = b$, that is $\lim_{x \rightarrow b^-} f(x) = \infty$. Then we define

$$\int_a^b f(x) dx = \lim_{t \rightarrow b^-} \int_a^t f(x) dx.$$

Now, if $f(x)$ is unbounded at an interior point $a < c < b$ then we define

$$\int_a^b f(x) dx = \lim_{t \rightarrow c^-} \int_a^t f(x) dx + \lim_{t \rightarrow c^+} \int_t^b f(x) dx.$$

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 42.6

Show that the improper integral $\int_0^1 \frac{1}{\sqrt{x}} dx$ converges.

Solution.

Indeed,

$$\begin{aligned} \int_0^1 \frac{1}{\sqrt{x}} dx &= \lim_{t \rightarrow 0^+} \int_t^1 \frac{1}{\sqrt{x}} dx = \lim_{t \rightarrow 0^+} 2\sqrt{x} \Big|_t^1 \\ &= \lim_{t \rightarrow 0^+} (2 - 2\sqrt{t}) = 2. \end{aligned}$$

In terms of area, we pick an $t > 0$ as shown in Figure 42.2:

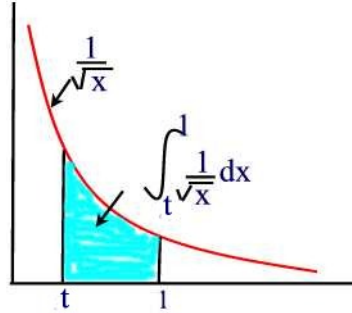


Figure 42.2

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. ■

Example 42.7

Investigate the convergence of $\int_0^2 \frac{1}{(x-2)^2} dx$.

Solution.

We deal with this improper integral as follows

$$\begin{aligned} \int_0^2 \frac{1}{(x-2)^2} dx &= \lim_{t \rightarrow 2^-} \int_0^t \frac{1}{(x-2)^2} dx = \lim_{t \rightarrow 2^-} -\frac{1}{(x-2)} \Big|_0^t \\ &= \lim_{t \rightarrow 2^-} \left(-\frac{1}{t-2} - \frac{1}{2}\right) = \infty. \end{aligned}$$

So that the given improper integral is divergent. ■

Example 42.8

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,

$$\begin{aligned} \int_{-1}^0 \frac{1}{x} dx &= \lim_{t \rightarrow 0^-} \int_{-1}^t \frac{1}{x} dx = \lim_{t \rightarrow 0^-} \ln |x| \Big|_{-1}^t \\ &= \lim_{t \rightarrow 0^-} \ln |t| = \infty. \end{aligned}$$

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 Types

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 42.9

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

$$\int_0^1 \frac{1}{x^2} dx = \lim_{t \rightarrow 0^+} \int_t^1 \frac{1}{x^2} dx = \lim_{t \rightarrow 0^+} -\frac{1}{x} \Big|_t^1 = \lim_{t \rightarrow 0^+} \left(\frac{1}{t} - 1 \right) = \infty.$$

Thus, $\int_0^1 \frac{1}{x^2} dx$ diverges and consequently the improper integral $\int_0^\infty \frac{1}{x^2} dx$ diverges. ■

43 Double Integrals

In this section, we introduce the concept of definite integral of a function of two variables over a rectangular region.

By a rectangular region we mean a region R as shown in Figure 43.1(I).

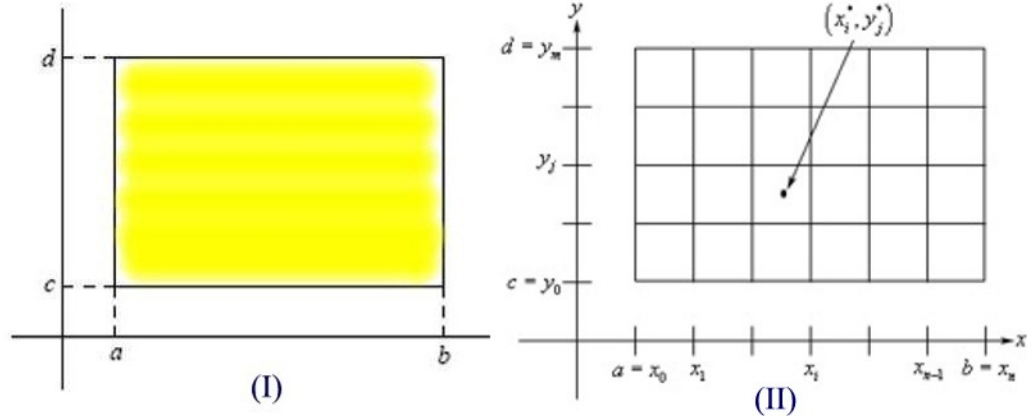


Figure 43.1

Let $f(x, y)$ be a continuous function on R . Our definition of the definite integral of f over the rectangle R will follow the definition from one variable calculus. Partition the interval $a \leq x \leq b$ into n equal subintervals using the mesh points $a \leq x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$ with $\Delta x = \frac{b-a}{n}$ denoting the length of each subinterval. Similarly, partition $c \leq y \leq d$ into m subintervals using the mesh points $c = y_0 < y_1 < y_2 < \cdots < y_{m-1} < y_m = d$ with $\Delta y = \frac{d-c}{m}$ denoting the length of each subinterval. This way, the rectangle R is partitioned into mn subrectangles each of area equals to $\Delta x \Delta y$ as shown in Figure 43.1(II).

Let D_{ij} be a typical rectangle. Let m_{ij} be the smallest value of f in D_{ij} and M_{ij} be the largest value in D_{ij} . Pick a point (x_i^*, y_j^*) in this rectangle. Then we can write

$$m_{ij} \Delta x \Delta y \leq f(x_i^*, y_j^*) \Delta x \Delta y \leq M_{ij} \Delta x \Delta y.$$

Sum over all i and j to obtain

$$\sum_{j=1}^m \sum_{i=1}^n m_{ij} \Delta x \Delta y \leq \sum_{j=1}^m \sum_{i=1}^n f(x_i^*, y_j^*) \Delta x \Delta y \leq \sum_{j=1}^m \sum_{i=1}^n M_{ij} \Delta x \Delta y.$$

We call

$$L = \sum_{j=1}^m \sum_{i=1}^n m_{ij} \Delta x \Delta y$$

the **lower Riemann sum** and

$$U = \sum_{j=1}^m \sum_{i=1}^n M_{ij} \Delta x \Delta y$$

the **upper Riemann sum**. If

$$\lim_{m,n \rightarrow \infty} L = \lim_{m,n \rightarrow \infty} U$$

then we write

$$\int_R f(x, y) dx dy = \lim_{m,n \rightarrow \infty} \sum_{j=1}^m \sum_{i=1}^n f(x_i^*, y_j^*) \Delta x \Delta y$$

and we call $\int_R f(x, y) dx dy$ the **double integral** of f over the rectangle R . The use of the word “double” will be justified below.

Double Integral as Volume Under a Surface

Just as the definite integral of a positive one-variable function can be interpreted as area, so the double integral of a positive two-variable function can be interpreted as a volume.

Let $f(x, y) > 0$ with surface S shown in Figure 43.2(I). Partition the rectangle R as above. Over each rectangle D_{ij} we will construct a box whose height is given by $f(x_i^*, y_j^*)$ as shown in Figure 43.2 (II). Each of the boxes has a base area of $\Delta x \Delta y$ and a height of $f(x_i^*, y_j^*)$ so the volume of each of these boxes is $f(x_i^*, y_j^*) \Delta x \Delta y$. So the volume under the surface S is then approximately,

$$V \approx \sum_{j=1}^m \sum_{i=1}^n f(x_i^*, y_j^*) \Delta x \Delta y$$

As the number of subdivisions grows, the tops of the boxes approximate the surface better, and the volume of the boxes gets closer and closer to the volume under the graph of the function. Thus, we have the following result,

If $f(x, y) > 0$ then the volume under the graph of f above the region R is

$$\int_R f(x, y) dx dy$$

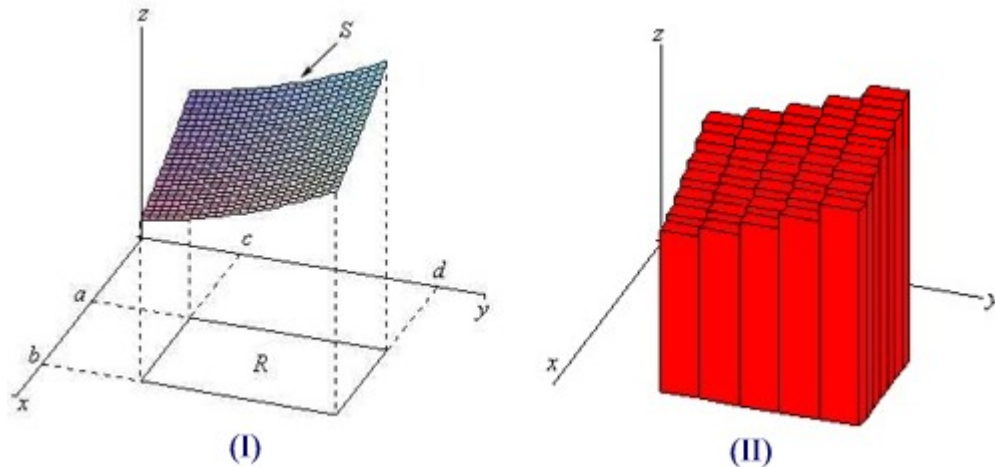


Figure 43.2

Double Integral as Area

If we choose a function $f(x, y) = 1$ everywhere in R then our integral becomes:

$$\text{Area of } R = \int_R 1 dx dy$$

That is, when $f(x, y) = 1$, the integral gives us the area of the region we are integrating over.

Example 43.1

Use the Riemann sum with $n = 3$, $m = 2$ and sample point the upper right corner of each subrectangle to estimate the volume under $z = xy$ and above the rectangle $0 \leq x \leq 6$, $0 \leq y \leq 4$.

Solution.

The interval on the x -axis is to be divided into $n = 3$ subintervals of equal length, so $\Delta x = \frac{6-0}{3} = 2$. Likewise, the interval on the y -axis is to be divided into $m = 2$ subintervals, also with width $\Delta y = 2$; and the rectangle is divided into six squares with sides 2.

Next, the upper right corners are at $(2, 2)$, $(4, 2)$ and $(6, 2)$, for the lower three squares, and $(2, 4)$, $(4, 4)$ and $(6, 4)$, for the upper three squares. The approximation is then

$$\begin{aligned} & [f(2, 2) + f(4, 2) + f(6, 2)] + [f(2, 4) + f(4, 4) + f(6, 4)] \cdot 2 \cdot 2 \\ & = [4 + 8 + 12 + 8 + 16 + 24] \cdot 4 = 72 \cdot 4 = 288 \blacksquare \end{aligned}$$

Example 43.2

Values of $f(x, y)$ are given in the table below. Let R be the rectangle $0 \leq x \leq 1.2$, $2 \leq y \leq 2.4$. Find Riemann sums which are reasonable over- and under-estimates for $\int_R f(x, y) dx dy$ with $\Delta x = 0.1$ and $\Delta y = 0.2$

$y \backslash x$	1.0	1.1	1.2
2.0	5	7	10
2.2	4	6	8
2.4	3	5	4

Solution.

We mark the values of the function on the plane, as shown in Figure 43.3, so that we can guess respectively at the smallest and largest value the function takes on each subrectangle.

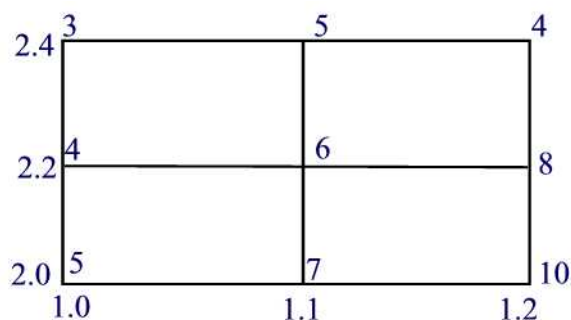


Figure 43.3

$$\text{Lower sum} = (4 + 6 + 3 + 4)\Delta x \Delta y = (17)(0.1)(0.2) = 0.34$$

$$\text{Upper sum} = (7 + 10 + 6 + 8)\Delta x \Delta y = (31)(0.1)(0.2) = 0.62 \blacksquare$$

Integral Over Bounded Regions That Are Not Rectangles

The region of integration R can be of any bounded shape not just rectangles. In our presentation above we chose a rectangular shaped region for convenience since it makes the summation limits and partitioning of the xy -plane into squares or rectangles simpler. However, this need not be the case. We can instead picture covering an arbitrary shaped region in the xy -plane with rectangles so that either all the rectangles lie just inside the region or the rectangles extend just outside the region (so that the region is contained inside our rectangles) as shown in Figure 43.4. We can then compute either the

minimum or maximum value of the function on each rectangle and compute the volume of the boxes, and sum.

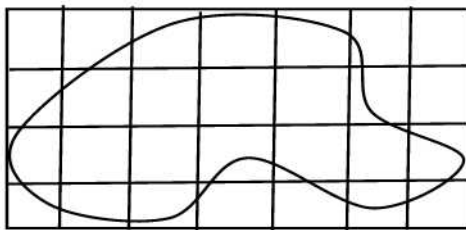


Figure 43.4

The Average of $f(x, y)$

As in the case of single variable calculus, the **average value** of $f(x, y)$ over a region R is defined by

$$\frac{1}{\text{Area of } R} \int_R f(x, y) dx dy.$$

Iterated Integrals

In the previous discussion we used Riemann sums to approximate a double integral. In this section, we see how to compute double integrals exactly using one-variable integrals.

Going back to our definition of the integral over a region as the limit of a double Riemann sum:

$$\begin{aligned} \int_R f(x, y) dx dy &= \lim_{m, n \rightarrow \infty} \sum_{j=1}^m \sum_{i=1}^n f(x_i^*, y_j^*) \Delta x \Delta y \\ &= \lim_{m, n \rightarrow \infty} \sum_{j=1}^m \left(\sum_{i=1}^n f(x_i^*, y_j^*) \Delta x \right) \Delta y \\ &= \lim_{m, n \rightarrow \infty} \sum_{j=1}^m \Delta y \left(\sum_{i=1}^n f(x_i^*, y_j^*) \Delta x \right) \\ &= \lim_{m \rightarrow \infty} \sum_{j=1}^m \Delta y \int_a^b f(x, y_j^*) dx \end{aligned}$$

We now let

$$F(y_j^*) = \int_a^b f(x, y_j^*) dx$$

and, substituting into the expression above, we obtain

$$\int_R f(x, y) dx dy = \lim_{m \rightarrow \infty} \sum_{j=1}^m F(y_j^*) \Delta y = \int_c^d F(y) dy = \int_c^d \int_a^b f(x, y) dx dy.$$

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 43.3

Compute $\int_0^{16} \int_0^8 (12 - \frac{x}{4} - \frac{y}{8}) dx dy$.

Solution.

We have

$$\begin{aligned} \int_0^{16} \int_0^8 \left(12 - \frac{x}{4} - \frac{y}{8}\right) dx dy &= \int_0^{16} \left(\int_0^8 \left(12 - \frac{x}{4} - \frac{y}{8}\right) dx \right) dy \\ &= \int_0^{16} \left[12x - \frac{x^2}{8} - \frac{xy}{8} \right]_0^8 dy \\ &= \int_0^{16} (88 - y) dy = 88y - \frac{y^2}{2} \Big|_0^{16} = 1280 \blacksquare \end{aligned}$$

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

$$\int_R f(x, y) dx dy = \int_a^b \int_c^d f(x, y) dy dx.$$

Example 43.4

Compute $\int_0^8 \int_0^{16} (12 - \frac{x}{4} - \frac{y}{8}) dy dx$.

Solution.

We have

$$\begin{aligned} \int_0^8 \int_0^{16} \left(12 - \frac{x}{4} - \frac{y}{8}\right) dy dx &= \int_0^8 \left(\int_0^{16} \left(12 - \frac{x}{4} - \frac{y}{8}\right) dy \right) dx \\ &= \int_0^8 \left[12y - \frac{xy}{4} - \frac{y^2}{16} \right]_0^{16} dx \\ &= \int_0^8 (176 - 4x) dx = 176x - 2x^2 \Big|_0^8 = 1280 \blacksquare \end{aligned}$$

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,

$$\int_R f(x, y) dx dy$$

where R is any region. We consider the two types of regions shown in Figure 43.5.

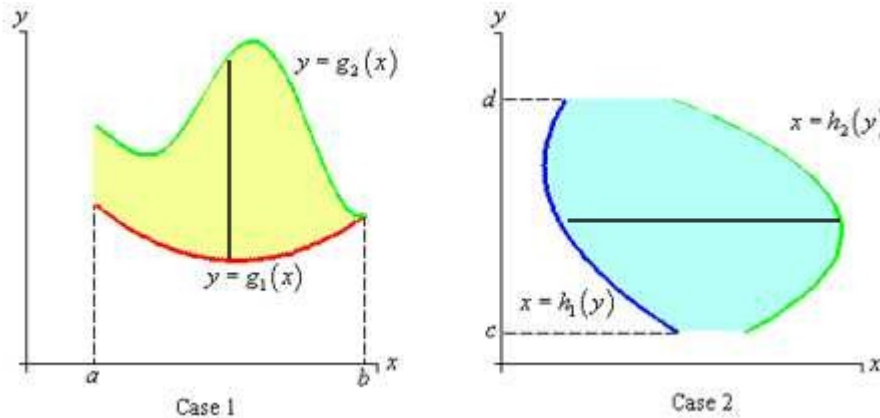


Figure 43.5

In Case 1, the iterated integral of f over R is defined by

$$\int_R f(x, y) dx dy = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) dy dx$$

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

$$\int_R f(x, y) dx dy = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y) dx dy$$

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 43.1

Choosing 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 43.5

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 43.6 shows the region of integration for this example.

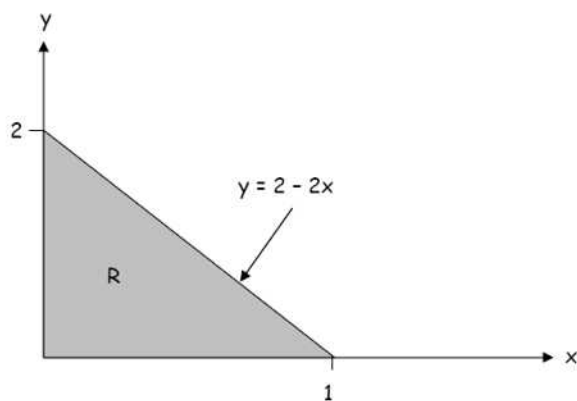


Figure 43.6

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 43.7(I).

$$\begin{aligned}\int_R xy dx dy &= \int_0^1 \int_0^{2-2x} xy dy dx \\ &= \int_0^1 \left. \frac{xy^2}{2} \right|_0^{2-2x} dx = \frac{1}{2} \int_0^1 x(2-2x)^2 dx \\ &= 2 \int_0^1 (x - 2x^2 + x^3) dx = 2 \left[\frac{x^2}{2} - \frac{2}{3}x^3 + \frac{x^4}{4} \right]_0^1 = \frac{1}{6}\end{aligned}$$

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 43.7(II)

$$\begin{aligned}\int_R xy dx dy &= \int_0^2 \int_0^{1-\frac{1}{2}y} xy dx dy \\ &= \int_0^2 \left. \frac{x^2 y}{2} \right|_0^{1-\frac{1}{2}y} dy = \frac{1}{2} \int_0^2 y \left(1 - \frac{1}{2}y\right)^2 dy \\ &= \frac{1}{2} \int_0^2 \left(y - y^2 + \frac{y^3}{4}\right) dy = \left. \frac{y^2}{4} - \frac{y^3}{6} + \frac{y^4}{32} \right|_0^2 = \frac{1}{6} \blacksquare\end{aligned}$$

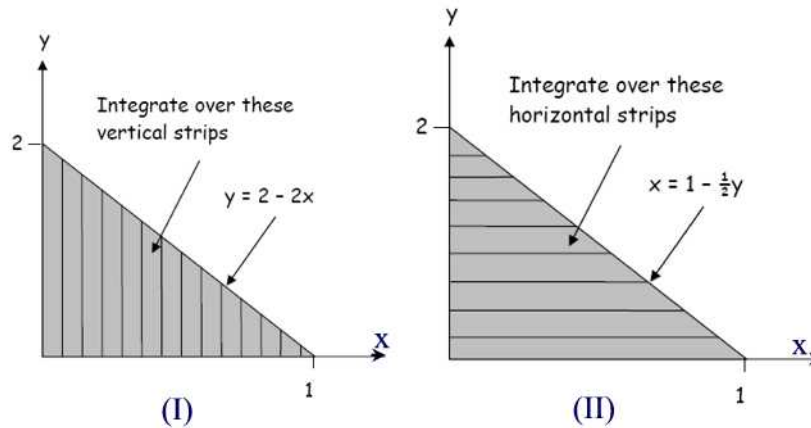


Figure 43.7

Example 43.6

Find $\int_R (4xy - y^3) dx dy$ 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 43.8. Using horizontal strips we can write

$$\begin{aligned} \int_R (4xy - y^3) dx dy &= \int_0^1 \int_{y^2}^{\sqrt[3]{y}} (4xy - y^3) dx dy \\ &= \int_0^1 2x^2 y - xy^3 \Big|_{y^2}^{\sqrt[3]{y}} dy = \int_0^1 \left(2y^{\frac{5}{3}} - y^{\frac{10}{3}} - y^5 \right) dy \\ &= \frac{3}{4} y^{\frac{8}{3}} - \frac{3}{13} y^{\frac{13}{3}} - \frac{1}{6} y^6 \Big|_0^1 = \frac{55}{156} \blacksquare \end{aligned}$$

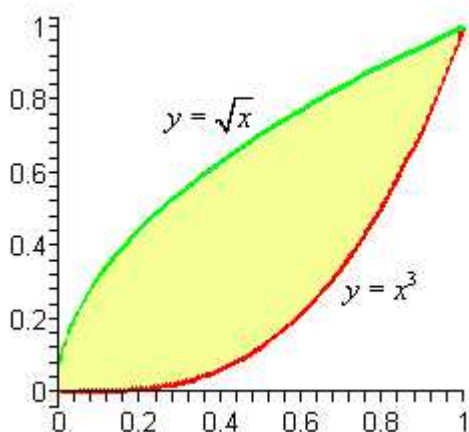


Figure 43.8

Example 43.7

Sketch the region of integration of $\int_0^2 \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} xy dy dx$

Solution.

A sketch of the region is given in Figure 43.9.

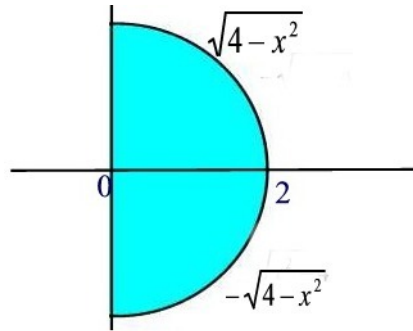


Figure 43.9

Problem 43.1

Set up a double integral of $f(x, y)$ over the region given by $0 < x < 1; x < y < x + 1$.

Problem 43.2

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 43.3

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 43.4

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 43.5

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 43.6

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 43.7

Evaluate $\int \int_R e^{-x-y} dx dy$, where R is the region in the first quadrant in which $x + y \leq 1$.

Problem 43.8

Evaluate $\int \int_R e^{-x-2y} dx dy$, where R is the region in the first quadrant in which $x \leq y$.

Problem 43.9

Evaluate $\int \int_R (x^2 + y^2) dx dy$, where R is the region $0 \leq x \leq y \leq L$.

Problem 43.10

Evaluate $\int \int_R f(x, y) dx dy$, where R is the region inside the unit square in which both coordinates x and y are greater than 0.5.

Problem 43.11

Evaluate $\int \int_R (x - y + 1) dx dy$, where R is the region inside the unit square in which $x + y \geq 0.5$.

Problem 43.12

Evaluate $\int_0^1 \int_0^1 x \max(x, y) dy dx$.

44 Double Integrals in Polar Coordinates

There are regions in the plane that are not easily used as domains of iterated integrals in rectangular coordinates. For instance, regions such as a disk, ring, or a portion of a disk or ring.

We start by recalling from Section 50 the relationship between Cartesian and polar coordinates.

The Cartesian system consists of two rectangular axes. A point P in this system is uniquely determined by two points x and y as shown in Figure 44.1(a). The polar coordinate system consists of a point O , called the **pole**, and a half-axis starting at O and pointing to the right, known as the **polar axis**. A point P in this system is determined by two numbers: the distance r between P and O and an angle θ between the ray OP and the polar axis as shown in Figure 44.1(b).

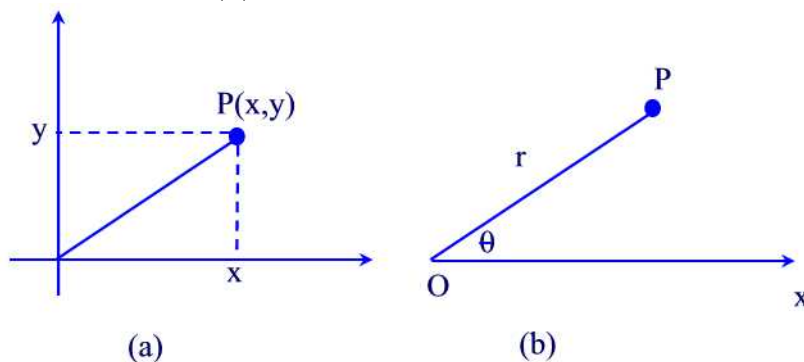


Figure 44.1

The Cartesian and polar coordinates can be combined into one figure as shown in Figure 44.2.

Figure 44.2 reveals the relationship between the Cartesian and polar coordinates:

$$r = \sqrt{x^2 + y^2} \quad x = r \cos \theta \quad y = r \sin \theta \quad \tan \theta = \frac{y}{x}.$$

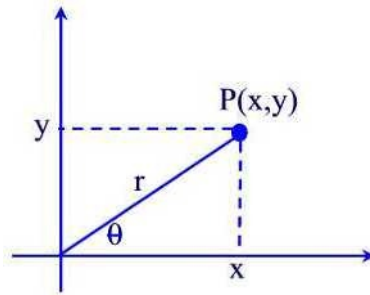


Figure 44.2

A double integral in polar coordinates can be defined as follows. Suppose we have a region

$$R = \{(r, \theta) : a \leq r \leq b, \alpha \leq \theta \leq \beta\}$$

as shown in Figure 44.3(a).

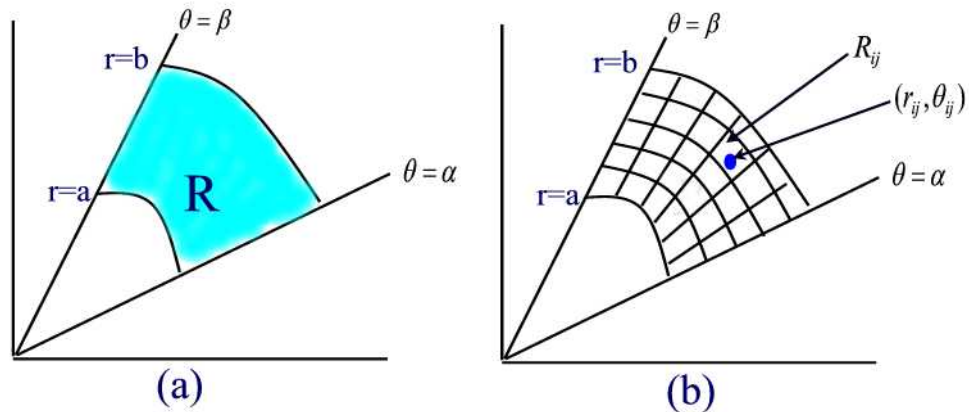


Figure 44.3

Partition the interval $\alpha \leq \theta \leq \beta$ into m equal subintervals, and the interval $a \leq r \leq b$ into n equal subintervals, thus obtaining mn subrectangles as shown in Figure 44.3(b). Choose a sample point (r_{ij}, θ_{ij}) in the subrectangle R_{ij} defined by $r_{i-1} \leq r \leq r_i$ and $\theta_{j-1} \leq \theta \leq \theta_j$. Then

$$\int_R f(x, y) dx dy \approx \sum_{j=1}^m \sum_{i=1}^n f(r_{ij}, \theta_{ij}) \Delta R_{ij}$$

where ΔR_{ij} is the area of the subrectangle R_{ij} .

To calculate the area of R_{ij} , look at Figure 44.4. If Δr and $\Delta\theta$ are small then R_{ij} is approximately a rectangle with area $r_{ij}\Delta r\Delta\theta$ so

$$\Delta R_{ij} \approx r_{ij}\Delta r\Delta\theta.$$

Thus, the double integral can be approximated by a Riemann sum

$$\int_R f(x, y) dx dy \approx \sum_{j=1}^m \sum_{i=1}^n f(r_{ij}, \theta_{ij}) r_{ij} \Delta r \Delta \theta$$

Taking the limit as $m, n \rightarrow \infty$ we obtain

$$\int_R f(x, y) dx dy = \int_{\alpha}^{\beta} \int_a^b f(r, \theta) r dr d\theta.$$

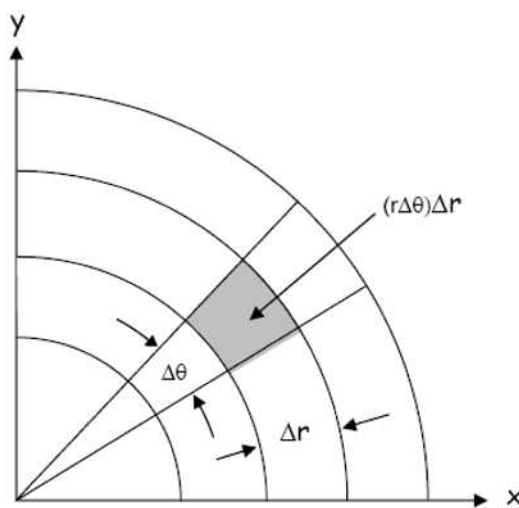


Figure 44.4

Example 44.1

Evaluate $\int_R e^{x^2+y^2} dx dy$ where $R : x^2 + y^2 \leq 1$.

Solution.

We have

$$\begin{aligned}\int_R e^{x^2+y^2} dx dy &= \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} e^{x^2+y^2} dy dx \\ &= \int_0^{2\pi} \int_0^1 e^{r^2} r dr d\theta = \int_0^{2\pi} \left. \frac{1}{2} e^{r^2} \right|_0^1 d\theta \\ &= \int_0^{2\pi} \frac{1}{2} (e - 1) d\theta = \pi(e - 1) \blacksquare\end{aligned}$$

Example 44.2

Compute the area of a circle of radius 1.

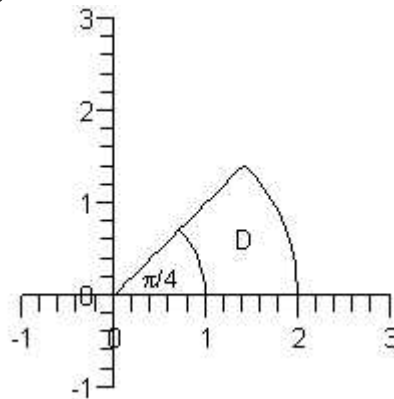
Solution.

The area is given by

$$\begin{aligned}\int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy dx &= \int_0^{2\pi} \int_0^1 r dr d\theta \\ &= \frac{1}{2} \int_0^{2\pi} d\theta = \pi \blacksquare\end{aligned}$$

Example 44.3

Evaluate $f(x, y) = \frac{1}{x^2+y^2}$ over the region D shown below.

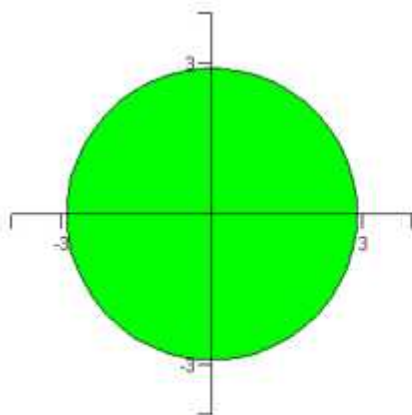
**Solution.**

We have

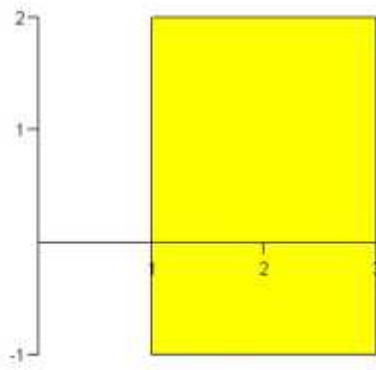
$$\int_0^{\pi/4} \int_1^2 \frac{1}{r^2} r dr d\theta = \int_0^{\pi/4} \ln 2 d\theta = \frac{\pi}{4} \ln 2 \blacksquare$$

Example 44.4

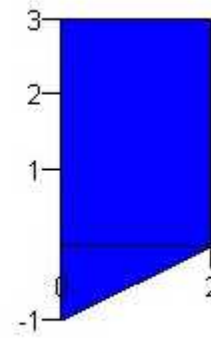
For each of the regions shown below, decide whether to integrate using rectangular or polar coordinates. In each case write an iterated integral of an arbitrary function $f(x, y)$ over the region.



(a)



(b)



(c)

Solution.

$$(a) \int_0^{2\pi} \int_0^3 f(r, \theta) r dr d\theta$$

$$(b) \int_1^3 \int_{-1}^2 f(x, y) dy dx$$

$$(c) \int_0^2 \int_{\frac{1}{2}x-1}^3 f(x, y) dy dx \blacksquare$$

Bibliography

- [1] R. Sheldon , *A first course in Probability*, 6th Edition (2002), Prentice Hall.
- [2] SOA/CAS, *Exam P Sample Questions*