Introduction to

Mathematical Statistics

Eighth Edition

Hogg

McKean

Craig

This page intentionally left blank

Introduction to Mathematical Statistics

Eighth Edition

Robert V. Hogg
University of Iowa

Joseph W. McKean
Western Michigan University

Allen T. Craig
Late Professor of Statistics
University of Iowa

Pearson

Director, Portfolio Management: Deirdre Lynch
Courseware Portfolio Manager: Patrick Barbera
Portfolio Management Assistant: Morgan Danna
Content Producer: Lauren Morse
Managing Producer: Scott Disanno
Product Marketing Manager: Yvonne Vannatta
Field Marketing Manager: Evan St. Cyr
Marketing Assistant: Jon Bryant
Senior Author Support/Technology Specialist: Joe Vetere
Manager, Rights and Permissions: Gina Cheselka
Manufacturing Buyer: Carol Melville, LSC Communications
Art Director: Barbara Atkinson
Production Coordination and Illustrations: Integra
Cover Design: Studio Montage
Cover Image: Aleksandarvelasevic/Digital Vision Vectors/Getty Images.

Library of Congress Cataloging-in-Publications Data
Names: Hogg, Robert V., author. | McKean, Joseph W., 1944- author. | Craig, Allen T. (Allen Thornton), 1905- author.
Title: Introduction to mathematical statistics / Robert V. Hogg, Late Professor of Statistics, University of Iowa, Joseph W. McKean, Western Michigan University, Allen T. Craig, Late Professor of Statistics, University of Iowa.
Description: Eighth edition. | Boston : Pearson, [2019] | Includes bibliographical references and index.
Identifiers: LCCN 2017033015| ISBN 9780134686998 | ISBN 0134686993
Subjects: LCSH: Mathematical statistics.
Classification: LCC QA276 .H59 2019 | DDC 519.5–dc23 LC record available at https://lccn.loc.gov/2017033015

ISBN 13: 978-0-13-468699-8

Pearson

ISBN 10: 0-13-468699-3

Dedicated to my wife Marge
and to the memory of Bob Hogg

This page intentionally left blank

Contents

This page intentionally left blank

Preface

We have made substantial changes in this edition of *Introduction to Mathematical Statistics*. Some of these changes help students appreciate the connection between statistical theory and statistical practice while other changes enhance the development and discussion of the statistical theory presented in this book.

Many of the changes in this edition reflect comments made by our readers. One of these comments concerned the small number of real data sets in the previous editions. In this edition, we have included more real data sets, using them to illustrate statistical methods or to compare methods. Further, we have made these data sets accessible to students by including them in the free R package `hmcpkg`. They can also be individually downloaded in an R session at the url listed below. In general, the R code for the analyses on these data sets is given in the text.

We have also expanded the use of the statistical software R. We selected R because it is a powerful statistical language that is free and runs on all three main platforms (Windows, Mac, and Linux). Instructors, though, can select another statistical package. We have also expanded our use of R functions to compute analyses and simulation studies, including several games. We have kept the level of coding for these functions straightforward. Our goal is to show students that with a few simple lines of code they can perform significant computations. Appendix B contains a brief R primer, which suffices for the understanding of the R used in the text. As with the data sets, these R functions can be sourced individually at the cited url; however, they are also included in the package `hmcpkg`.

We have supplemented the mathematical review material in Appendix A, placing it in the document *Mathematical Primer for Introduction to Mathematical Statistics*. It is freely available for students to download at the listed url. Besides sequences, this supplement reviews the topics of infinite series, differentiation, and integration (univariate and bivariate). We have also expanded the discussion of iterated integrals in the text. We have added figures to clarify discussion.

We have retained the order of elementary statistical inferences (Chapter 4) and asymptotic theory (Chapter 5). In Chapters 5 and 6, we have written brief reviews of the material in Chapter 4, so that Chapters 4 and 5 are essentially independent of one another and, hence, can be interchanged. In Chapter 3, we now begin the section on the multivariate normal distribution with a subsection on the bivariate normal distribution. Several important topics have been added. This includes Tukey's multiple comparison procedure in Chapter 9 and confidence intervals for the correlation coefficients found in Chapters 9 and 10. Chapter 7 now contains a discussion on standard errors for estimates obtained by bootstrapping the sample. Several topics that were discussed in the Exercises are now discussed in the text. Examples include quantiles, Section 1.7.1, and hazard functions, Section 3.3. In general, we have made more use of subsections to break up some of the discussion. Also, several more sections are now indicated by * as being optional.

Content and Course Planning

Chapters 1 and 2 develop probability models for univariate and multivariate variables while Chapter 3 discusses many of the most widely used probability models. Chapter 4 discusses statistical theory for much of the inference found in a standard statistical methods course. Chapter 5 presents asymptotic theory, concluding with the Central Limit Theorem. Chapter 6 provides a complete inference (estimation and testing) based on maximum likelihood theory. The EM algorithm is also discussed. Chapters 7–8 contain optimal estimation procedures and tests of statistical hypotheses. The final three chapters provide theory for three important topics in statistics. Chapter 9 contains inference for normal theory methods for basic analysis of variance, univariate regression, and correlation models. Chapter 10 presents nonparametric methods (estimation and testing) for location and univariate regression models. It also includes discussion on the robust concepts of efficiency, influence, and breakdown. Chapter 11 offers an introduction to Bayesian methods. This includes traditional Bayesian procedures as well as Markov Chain Monte Carlo techniques.

Several courses can be designed using our book. The basic two-semester course in mathematical statistics covers most of the material in Chapters 1–8 with topics selected from the remaining chapters. For such a course, the instructor would have the option of interchanging the order of Chapters 4 and 5, thus beginning the second semester with an introduction to statistical theory (Chapter 4). A one-semester course could consist of Chapters 1–4 with a selection of topics from Chapter 5. Under this option, the student sees much of the statistical theory for the methods discussed in a non-theoretical course in methods. On the other hand, as with the two-semester sequence, after covering Chapters 1–3, the instructor can elect to cover Chapter 5 and finish the course with a selection of topics from Chapter 4.

The data sets and R functions used in this book and the R package `hmcpkg` can be downloaded at the site:
`https://media.pearsoncmg.com/cmg/pmmg_mml_shared/mathstatsresources/home/index.html`

Acknowledgements

Bob Hogg passed away in 2014, so he did not work on this edition of the book. Often, though, when I was trying to decide whether or not to make a change in the manuscript, I found myself thinking of what Bob would do. In his memory, I have retained the order of the authors for this edition.

As with earlier editions, comments from readers are always welcomed and appreciated. We would like to thank these reviewers of the previous edition: James Baldone, Virginia College; Steven Culpepper, University of Illinois at Urbana-Champaign; Yuichiro Kakihara, California State University; Jaechoul Lee, Boise State University; Michael Levine, Purdue University; Tingni Sun, University of Maryland, College Park; and Daniel Weiner, Boston University. We appreciated and took into consideration their comments for this revision. We appreciate the helpful comments of Thomas Hettmansperger of Penn State University, Ash Abebe of Auburn University, and Professor Ioannis Kalogridis of the University of Leuven. A special thanks to Patrick Barbera (Portfolio Manager, Statistics), Lauren Morse (Content Producer, Math/Stats), Yvonne Vannatta (Product Marketing Manager), and the rest of the staff at Pearson for their help in putting this edition together. Thanks also to Richard Ponticelli, North Shore Community College, who accuracy checked the page proofs. Also, a special thanks to my wife Marge for her unwavering support and encouragement of my efforts in writing this edition.

Joe McKean

This page intentionally left blank

Chapter 1

Probability and Distributions

1.1 Introduction

In this section, we intuitively discuss the concepts of a probability model which we formalize in Secton 1.3 Many kinds of investigations may be characterized in part by the fact that repeated experimentation, under essentially the same conditions, is more or less standard procedure. For instance, in medical research, interest may center on the effect of a drug that is to be administered; or an economist may be concerned with the prices of three specified commodities at various time intervals; or an agronomist may wish to study the effect that a chemical fertilizer has on the yield of a cereal grain. The only way in which an investigator can elicit information about any such phenomenon is to perform the experiment. Each experiment terminates with an *outcome*. But it is characteristic of these experiments that the outcome cannot be predicted with certainty prior to the experiment.

Suppose that we have such an experiment, but the experiment is of such a nature that a collection of every possible outcome can be described prior to its performance. If this kind of experiment can be repeated under the same conditions, it is called a **random experiment**, and the collection of every possible outcome is called the experimental space or the **sample space**. We denote the sample space by $\mathcal{C}$.

Example 1.1.1. In the toss of a coin, let the outcome tails be denoted by T and let the outcome heads be denoted by H. If we assume that the coin may be repeatedly tossed under the same conditions, then the toss of this coin is an example of a random experiment in which the outcome is one of the two symbols T or H; that is, the sample space is the collection of these two symbols. For this example, then, $\mathcal{C} = \{H, T\}$. ■

Example 1.1.2. In the cast of one red die and one white die, let the outcome be the ordered pair (number of spots up on the red die, number of spots up on the white die). If we assume that these two dice may be repeatedly cast under the same conditions, then the cast of this pair of dice is a random experiment. The sample space consists of the 36 ordered pairs: $\mathcal{C} = \{(1,1), \ldots, (1,6), (2,1), \ldots, (2,6), \ldots, (6,6)\}$. ■

We generally use small Roman letters for the elements of $\mathcal{C}$ such as a, b, or c. Often for an experiment, we are interested in the chances of certain subsets of elements of the sample space occurring. Subsets of $\mathcal{C}$ are often called **events** and are generally denoted by capitol Roman letters such as $A, B,$ or C. If the experiment results in an element in an event A, we say the event A has occurred. We are interested in the chances that an event occurs. For instance, in Example 1.1.1 we may be interested in the chances of getting heads; i.e., the chances of the event $A = \{H\}$ occurring. In the second example, we may be interested in the occurrence of the sum of the upfaces of the dice being "7" or "11;" that is, in the occurrence of the event $A = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1),(5,6),(6,5)\}$.

Now conceive of our having made N repeated performances of the random experiment. Then we can count the number f of times (the **frequency**) that the event A actually occurred throughout the N performances. The ratio f/N is called the **relative frequency** of the event A in these N experiments. A relative frequency is usually quite erratic for small values of N, as you can discover by tossing a coin. But as N increases, experience indicates that we associate with the event A a number, say p, that is equal or approximately equal to that number about which the relative frequency seems to stabilize. If we do this, then the number p can be interpreted as that number which, in future performances of the experiment, the relative frequency of the event A will either equal or approximate. Thus, although we *cannot* predict the outcome of a random experiment, we *can*, for a large value of N, predict approximately the relative frequency with which the outcome will be in A. The number p associated with the event A is given various names. Sometimes it is called the *probability* that the outcome of the random experiment is in A; sometimes it is called the *probability* of the event A; and sometimes it is called the *probability measure* of A. The context usually suggests an appropriate choice of terminology.

Example 1.1.3. Let $\mathcal{C}$ denote the sample space of Example 1.1.2 and let B be the collection of every ordered pair of $\mathcal{C}$ for which the sum of the pair is equal to seven. Thus $B = \{(1,6),(2,5),(3,4),(4,3),(5,2)(6,1)\}$. Suppose that the dice are cast $N = 400$ times and let f denote the frequency of a sum of seven. Suppose that 400 casts result in $f = 60$. Then the relative frequency with which the outcome was in B is $f/N = \frac{60}{400} = 0.15$. Thus we might associate with B a number p that is close to 0.15, and p would be called the probability of the event B. ■

Remark 1.1.1. The preceding interpretation of probability is sometimes referred to as the *relative frequency approach*, and it obviously depends upon the fact that an experiment can be repeated under essentially identical conditions. However, many persons extend probability to other situations by treating it as a rational measure of belief. For example, the statement $p = \frac{2}{5}$ for an event A would mean to them that their *personal* or *subjective* probability of the event A is equal to $\frac{2}{5}$. Hence, if they are not opposed to gambling, this could be interpreted as a willingness on their part to bet on the outcome of A so that the two possible payoffs are in the ratio $p/(1-p) = \frac{2}{5}/\frac{3}{5} = \frac{2}{3}$. Moreover, if they truly believe that $p = \frac{2}{5}$ is correct, they would be willing to accept either side of the bet: (a) win 3 units if A occurs and lose 2 if it does not occur, or (b) win 2 units if A does not occur and lose 3 if it does. However, since the mathematical properties of probability given in Section 1.3 are consistent with either of these interpretations, the subsequent mathematical development does not depend upon which approach is used. ■

The primary purpose of having a mathematical theory of statistics is to provide mathematical models for random experiments. Once a model for such an experiment has been provided and the theory worked out in detail, the statistician may, within this framework, make inferences (that is, draw conclusions) about the random experiment. The construction of such a model requires a theory of probability. One of the more logically satisfying theories of probability is that based on the concepts of sets and functions of sets. These concepts are introduced in Section 1.2.

1.2 Sets

The concept of a *set* or a *collection* of objects is usually left undefined. However, a particular set can be described so that there is no misunderstanding as to what collection of objects is under consideration. For example, the set of the first 10 positive integers is sufficiently well described to make clear that the numbers $\frac{3}{4}$ and 14 are not in the set, while the number 3 is in the set. If an object belongs to a set, it is said to be an *element* of the set. For example, if C denotes the set of real numbers x for which $0 \leq x \leq 1$, then $\frac{3}{4}$ is an element of the set C. The fact that $\frac{3}{4}$ is an element of the set C is indicated by writing $\frac{3}{4} \in C$. More generally, $c \in C$ means that c is an element of the set C.

The sets that concern us are frequently *sets of numbers*. However, the language of sets of *points* proves somewhat more convenient than that of sets of numbers. Accordingly, we briefly indicate how we use this terminology. In analytic geometry considerable emphasis is placed on the fact that to each point on a line (on which an origin and a unit point have been selected) there corresponds one and only one number, say x; and that to each number x there corresponds one and only one point on the line. This one-to-one correspondence between the numbers and points on a line enables us to speak, without misunderstanding, of the "point x" instead of the "number x." Furthermore, with a plane rectangular coordinate system and with x and y numbers, to each symbol (x, y) there corresponds one and only one point in the plane; and to each point in the plane there corresponds but one such symbol. Here again, we may speak of the "point (x, y)," meaning the "ordered number pair x and y." This convenient language can be used when we have a rectangular coordinate system in a space of three or more dimensions. Thus the "point $(x_1, x_2, \ldots, x_n)$" means the numbers $x_1, x_2, \ldots, x_n$ in the order stated. Accordingly, in describing our sets, we frequently speak of a set of points (a set whose elements are points), being careful, of course, to describe the set so as to avoid any ambiguity. The notation $C = \{x : 0 \leq x \leq 1\}$ is read "C is the one-dimensional set of points x for which $0 \leq x \leq 1$." Similarly, $C = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$ can be read "C is the two-dimensional set of points (x, y) that are interior to, or on the boundary of, a square with opposite vertices at $(0, 0)$ and $(1, 1)$."

We say a set C is **countable** if C is finite or has as many elements as there are positive integers. For example, the sets $C_1 = \{1, 2, \ldots, 100\}$ and $C_2 = \{1, 3, 5, 7, \ldots\}$

are countable sets. The interval of real numbers $(0, 1]$, though, is not countable.

1.2.1 Review of Set Theory

As in Section 1.1, let $\mathcal{C}$ denote the sample space for the experiment. Recall that events are subsets of $\mathcal{C}$. We use the words event and subset interchangeably in this section. An elementary algebra of sets will prove quite useful for our purposes. We now review this algebra below along with illustrative examples. For illustration, we also make use of **Venn diagrams**. Consider the collection of Venn diagrams in Figure 1.2.1. The interior of the rectangle in each plot represents the sample space C. The shaded region in Panel (a) represents the event A.

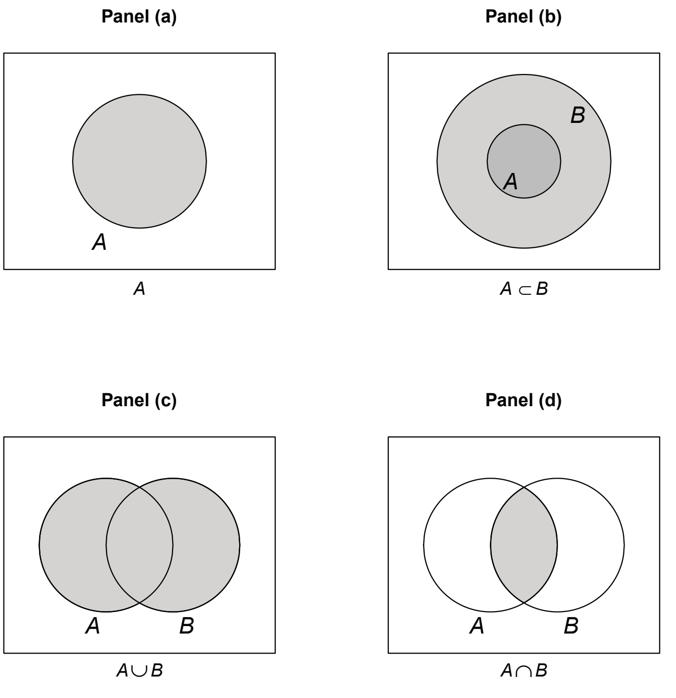

Figure 1.2.1: A series of Venn diagrams. The sample space C is represented by the interior of the rectangle in each plot. Panel (a) depicts the event A; Panel (b) depicts $A \subset B$; Panel (c) depicts $A \cup B$; and Panel (d) depicts $A \cap B$.

We first define the complement of an event A.

Definition 1.2.1. *The* **complement** *of an event A is the set of all elements in C which are not in A. We denote the complement of A by A^c. That is, $A^c = \{x \in \mathcal{C} : x \notin A\}$.*

The complement of A is represented by the white space in the Venn diagram in Panel (a) of Figure 1.2.1.

The empty set is the event with no elements in it. It is denoted by ϕ. Note that $\mathcal{C}^c = \phi$ and $\phi^c = \mathcal{C}$. The next definition defines when one event is a subset of another.

Definition 1.2.2. *If each element of a set A is also an element of set B, the set A is called a* **subset** *of the set B. This is indicated by writing $A \subset B$. If $A \subset B$ and also $B \subset A$, the two sets have the same elements, and this is indicated by writing $A = B$.*

Panel (b) of Figure 1.2.1 depicts $A \subset B$.

The event A or B is defined as follows:

Definition 1.2.3. *Let A and B be events. Then the* **union** *of A and B is the set of all elements that are in A or in B or in both A and B. The union of A and B is denoted by $A \cup B$*

Panel (c) of Figure 1.2.1 shows $A \cup B$.

The event that both A and B occur is defined by,

Definition 1.2.4. *Let A and B be events. Then the* **intersection** *of A and B is the set of all elements that are in both A and B. The intersection of A and B is denoted by $A \cap B$*

Panel (d) of Figure 1.2.1 illustrates $A \cap B$.

Two events are **disjoint** if they have no elements in common. More formally we define

Definition 1.2.5. *Let A and B be events. Then A and B are* **disjoint** *if $A \cap B = \phi$*

If A and B are disjoint, then we say $A \cup B$ forms a **disjoint union.** The next two examples illustrate these concepts.

Example 1.2.1. Suppose we have a spinner with the numbers 1 through 10 on it. The experiment is to spin the spinner and record the number spun. Then $\mathcal{C} = \{1, 2, \ldots, 10\}$. Define the events A, B, and C by $A = \{1, 2\}$, $B = \{2, 3, 4\}$, and $C = \{3, 4, 5, 6\}$, respectively.

$$A^c = \{3, 4, \ldots, 10\}; \quad A \cup B = \{1, 2, 3, 4\}; \quad A \cap B = \{2\}$$
$$A \cap C = \phi; \quad B \cap C = \{3, 4\}; \quad B \cap C \subset B; \quad B \cap C \subset C$$
$$A \cup (B \cap C) = \{1, 2\} \cup \{3, 4\} = \{1, 2, 3, 4\} \tag{1.2.1}$$
$$(A \cup B) \cap (A \cup C) = \{1, 2, 3, 4\} \cap \{1, 2, 3, 4, 5, 6\} = \{1, 2, 3, 4\} \tag{1.2.2}$$

The reader should verify these results. ■

Example 1.2.2. For this example, suppose the experiment is to select a real number in the open interval $(0, 5)$; hence, the sample space is $\mathcal{C} = (0, 5)$. Let $A = (1, 3)$,

$B = (2, 4)$, and $C = [3, 4.5)$.

$$A \cup B = (1, 4); \quad A \cap B = (2, 3); \quad B \cap C = [3, 4)$$

$$A \cap (B \cup C) = (1, 3) \cap (2, 4.5) = (2, 3) \tag{1.2.3}$$

$$(A \cap B) \cup (A \cap C) = (2, 3) \cup \phi = (2, 3) \tag{1.2.4}$$

A sketch of the real number line between 0 and 5 helps to verify these results. ■

Expressions (1.2.1)–(1.2.2) and (1.2.3)–(1.2.4) are illustrations of general **distributive laws**. For any sets A, B, and C,

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned} \tag{1.2.5}$$

These follow directly from set theory. To verify each identity, sketch Venn diagrams of both sides.

The next two identities are collectively known as **DeMorgan's Laws**. For any sets A and B,

$$(A \cap B)^c = A^c \cup B^c \tag{1.2.6}$$

$$(A \cup B)^c = A^c \cap B^c. \tag{1.2.7}$$

For instance, in Example 1.2.1,

$$(A \cup B)^c = \{1, 2, 3, 4\}^c = \{5, 6, \ldots, 10\} = \{3, 4, \ldots, 10\} \cap \{\{1, 5, 6, \ldots, 10\} = A^c \cap B^c;$$

while, from Example 1.2.2,

$$(A \cap B)^c = (2, 3)^c = (0, 2] \cup [3, 5) = [(0, 1] \cup [3, 5)] \cup [(0, 2] \cup [4, 5)] = A^c \cup B^c.$$

As the last expression suggests, it is easy to extend unions and intersections to more than two sets. If $A_1, A_2, \ldots, A_n$ are any sets, we define

$$A_1 \cup A_2 \cup \cdots \cup A_n = \{x : x \in A_i, \text{ for some } i = 1, 2, \ldots, n\} \tag{1.2.8}$$

$$A_1 \cap A_2 \cap \cdots \cap A_n = \{x : x \in A_i, \text{ for all } i = 1, 2, \ldots, n\}. \tag{1.2.9}$$

We often abbreviative these by $\cup_{i=1}^n A_i$ and $\cap_{i=1}^n A_i$, respectively. Expressions for countable unions and intersections follow directly; that is, if $A_1, A_2, \ldots, A_n \ldots$ is a sequence of sets then

$$A_1 \cup A_2 \cup \cdots = \{x : x \in A_n, \text{ for some } n = 1, 2, \ldots\} = \cup_{n=1}^{\infty} A_n \tag{1.2.10}$$

$$A_1 \cap A_2 \cap \cdots = \{x : x \in A_n, \text{ for all } n = 1, 2, \ldots\} = \cap_{n=1}^{\infty} A_n. \tag{1.2.11}$$

The next two examples illustrate these ideas.

Example 1.2.3. Suppose $\mathcal{C} = \{1, 2, 3, \ldots\}$. If $A_n = \{1, 3, \ldots, 2n-1\}$ and $B_n = \{n, n+1, \ldots\}$, for $n = 1, 2, 3, \ldots$, then

$$\cup_{n=1}^{\infty} A_n = \{1, 3, 5, \ldots\}; \quad \cap_{n=1}^{\infty} A_n = \{1\}; \tag{1.2.12}$$

$$\cup_{n=1}^{\infty} B_n = \mathcal{C}; \quad \cap_{n=1}^{\infty} B_n = \phi. \quad ■ \tag{1.2.13}$$

Example 1.2.4. Suppose $\mathcal{C}$ is the interval of real numbers $(0,5)$. Suppose $C_n = (1-n^{-1}, 2+n^{-1})$ and $D_n = (n^{-1}, 3-n^{-1})$, for $n = 1, 2, 3, \ldots$. Then

$$\cup_{n=1}^{\infty} C_n = (0,3); \quad \cap_{n=1}^{\infty} C_n = [1,2] \tag{1.2.14}$$

$$\cup_{n=1}^{\infty} D_n = (0,3); \quad \cap_{n=1}^{\infty} D_n = (1,2). \ \blacksquare \tag{1.2.15}$$

We occassionally have sequences of sets that are **monotone**. They are of two types. We say a sequence of sets $\{A_n\}$ is **nondecreasing, (nested upward)**, if $A_n \subset A_{n+1}$ for $n = 1, 2, 3, \ldots$. For such a sequence, we define

$$\lim_{n\to\infty} A_n = \cup_{n=1}^{\infty} A_n. \tag{1.2.16}$$

The sequence of sets $A_n = \{1, 3, \ldots, 2n-1\}$ of Example 1.2.3 is such a sequence. So in this case, we write $\lim_{n\to\infty} A_n = \{1, 3, 5, \ldots\}$. The sequence of sets $\{D_n\}$ of Example 1.2.4 is also a nondecreasing suquence of sets.

The second type of monotone sets consists of the **nonincreasing, (nested downward)** sequences. A sequence of sets $\{A_n\}$ is **nonincreasing**, if $A_n \supset A_{n+1}$ for $n = 1, 2, 3, \ldots$. In this case, we define

$$\lim_{n\to\infty} A_n = \cap_{n=1}^{\infty} A_n. \tag{1.2.17}$$

The sequences of sets $\{B_n\}$ and $\{C_n\}$ of Examples 1.2.3 and 1.2.4, respectively, are examples of nonincreasing sequences of sets.

1.2.2 Set Functions

Many of the functions used in calculus and in this book are functions that map real numbers into real numbers. We are concerned also with functions that map sets into real numbers. Such functions are naturally called functions of a set or, more simply, **set functions**. Next we give some examples of set functions and evaluate them for certain simple sets.

Example 1.2.5. Let $\mathcal{C} = R$, the set of real numbers. For a subset A in $\mathcal{C}$, let $Q(A)$ be equal to the number of points in A that correspond to positive integers. Then $Q(A)$ is a set function of the set A. Thus, if $A = \{x : 0 < x < 5\}$, then $Q(A) = 4$; if $A = \{-2, -1\}$, then $Q(A) = 0$; and if $A = \{x : -\infty < x < 6\}$, then $Q(A) = 5$. $\blacksquare$

Example 1.2.6. Let $\mathcal{C} = R^2$. For a subset A of $\mathcal{C}$, let $Q(A)$ be the area of A if A has a finite area; otherwise, let $Q(A)$ be undefined. Thus, if $A = \{(x,y) : x^2 + y^2 \le 1\}$, then $Q(A) = \pi$; if $A = \{(0,0), (1,1), (0,1)\}$, then $Q(A) = 0$; and if $A = \{(x,y) : 0 \le x, 0 \le y, x + y \le 1\}$, then $Q(A) = \frac{1}{2}$. $\blacksquare$

Often our set functions are defined in terms of sums or integrals.[1] With this in mind, we introduce the following notation. The symbol

$$\int_A f(x)\, dx$$

[1]Please see Chapters 2 and 3 of *Mathematical Comments*, at site noted in the Preface, for a review of sums and integrals

means the ordinary (Riemann) integral of $f(x)$ over a prescribed one-dimensional set A and the symbol

$$\iint_A g(x,y)\,dxdy$$

means the Riemann integral of $g(x,y)$ over a prescribed two-dimensional set A. This notation can be extended to integrals over n dimensions. To be sure, unless these sets A and these functions $f(x)$ and $g(x,y)$ are chosen with care, the integrals frequently fail to exist. Similarly, the symbol

$$\sum_A f(x)$$

means the sum extended over all $x \in A$ and the symbol

$$\sum_A\sum g(x,y)$$

means the sum extended over all $(x,y) \in A$. As with integration, this notation extends to sums over n dimensions.

The first example is for a set function defined on sums involving a **geometric series**. As pointed out in Example 2.3.1 of *Mathematical Comments*,[2] if $|a| < 1$, then the following series converges to $1/(1-a)$:

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}, \quad \text{if } |a| < 1. \tag{1.2.18}$$

Example 1.2.7. Let $\mathcal{C}$ be the set of all nonnegative integers and let A be a subset of $\mathcal{C}$. Define the set function Q by

$$Q(A) = \sum_{n\in A}\left(\frac{2}{3}\right)^n. \tag{1.2.19}$$

It follows from (1.2.18) that $Q(\mathcal{C}) = 3$. If $A = \{1,2,3\}$ then $Q(A) = 38/27$. Suppose $B = \{1,3,5,\ldots\}$ is the set of all odd positive integers. The computation of $Q(B)$ is given next. This derivation consists of rewriting the series so that (1.2.18) can be applied. Frequently, we perform such derivations in this book.

$$\begin{aligned} Q(B) &= \sum_{n\in B}\left(\frac{2}{3}\right)^n = \sum_{n=0}^{\infty}\left(\frac{2}{3}\right)^{2n+1} \\ &= \frac{2}{3}\sum_{n=0}^{\infty}\left[\left(\frac{2}{3}\right)^2\right]^n = \frac{2}{3}\frac{1}{1-(4/9)} = \frac{6}{5} \quad \blacksquare \end{aligned}$$

In the next example, the set function is defined in terms of an integral involving the exponential function $f(x) = e^{-x}$.

[2]Downloadable at site noted in the Preface

Example 1.2.8. Let $\mathcal{C}$ be the interval of positive real numbers, i.e., $\mathcal{C} = (0, \infty)$. Let A be a subset of $\mathcal{C}$. Define the set function Q by

$$Q(A) = \int_A e^{-x}\,dx, \tag{1.2.20}$$

provided the integral exists. The reader should work through the following integrations:

$$\begin{aligned}
Q[(1,3)] &= \int_1^3 e^{-x}\,dx = -e^{-x}\Big|_1^3 = e^{-1} - e^{-3} \doteq 0.318 \\
Q[(5,\infty)] &= \int_1^3 e^{-x}\,dx = -e^{-x}\Big|_5^\infty = e^{-5} \doteq 0.007 \\
Q[(1,3)\cup[3,5)] &= \int_1^5 e^{-x}\,dx = \int_1^3 e^{-x}\,dx + \int_3^5 e^{-x}\,dx = Q[(1,3)] + Q([3,5)] \\
Q(\mathcal{C}) &= \int_0^\infty e^{-x}\,dx = 1. \quad \blacksquare
\end{aligned}$$

Our final example, involves an n dimensional integral.

Example 1.2.9. Let $\mathcal{C} = R^n$. For A in $\mathcal{C}$ define the set function

$$Q(A) = \int \cdots \int_A dx_1 dx_2 \cdots dx_n,$$

provided the integral exists. For example, if $A = \{(x_1, x_2, \ldots, x_n) : 0 \le x_1 \le x_2, 0 \le x_i \le 1, \text{ for } 1 = 3, 4, \ldots, n\}$, then upon expressing the multiple integral as an iterated integral[3] we obtain

$$\begin{aligned}
Q(A) &= \int_0^1 \left[\int_0^{x_2} dx_1\right] dx_2 \bullet \prod_{i=3}^{n} \left[\int_0^1 dx_i\right] \\
&= \frac{x_2^2}{2}\Big|_0^1 \bullet 1 = \frac{1}{2}.
\end{aligned}$$

If $B = \{(x_1, x_2, \ldots, x_n) : 0 \le x_1 \le x_2 \le \cdots \le x_n \le 1\}$, then

$$\begin{aligned}
Q(B) &= \int_0^1 \left[\int_0^{x_n} \cdots \left[\int_0^{x_3} \left[\int_0^{x_2} dx_1\right] dx_2\right] \cdots dx_{n-1}\right] dx_n \\
&= \frac{1}{n!},
\end{aligned}$$

where $n! = n(n-1)\cdots 3 \cdot 2 \cdot 1$. $\blacksquare$

[3]For a discussion of multiple integrals in terms of iterated integrals, see Chapter 3 of *Mathematical Comments*.