

A

Math Review: Sets, Functions, Permutations, Combinations, and Notation



- A.1 Introduction
- A.2 Definitions, Axioms, Theorems, Corollaries, and Lemmas
- A.3 Elements of Set Theory
- A.4 Relations, Point Functions, and Set Functions
- A.5 Combinations and Permutations
- A.6 Summation, Integration and Matrix Differentiation Notation

A.1 Introduction

In this appendix we review basic results concerning set theory, relations and functions, combinations and permutations, and summation and integration notation. We also review the meaning of the terms *definition*, *axiom*, *theorem*, *corollary*, and *lemma*, which are labels that are affixed to a myriad of statements and results that constitute the theory of probability and mathematical statistics. The topics reviewed in this appendix constitute basic foundational material on which our study of mathematical statistics will be based. Additional mathematical results, often of a more advanced nature, are introduced throughout the text as the need arises.

A.2 Definitions, Axioms, Theorems, Corollaries, and Lemmas

The development of the theory of probability and mathematical statistics involves a considerable number of statements consisting of definitions, axioms, theorems, corollaries, and lemmas. These terms will be used for organizing the various statements and results we will examine into these categories:

1. descriptions of meaning;
2. statements that are acceptable as true without proof;

3. formulas or statements that require proof of validity;
4. formulas or statements whose validity follows immediately from other true formulas or statements; and
5. results, generally from other branches of mathematics, whose primary purpose is to facilitate the proof of validity of formulas or statements in mathematical statistics.

More formally, we present the following meaning of the terms:

Definition A statement of the meaning of a word, word group, sign, or symbol.

Axiom (or postulate) A statement that has found general acceptance, or is thought to be worthy thereof, on the basis of an appeal to intrinsic merit or self-evidence and thus requires no proof of validity.

Theorem (or proposition) A formula or statement that is deduced from other proved or accepted formulas or statements and whose validity is thereby proved.

Corollary A formula or statement that is immediately deducible from a proven theorem and that requires little or no additional proof of validity.

Lemma An auxiliary proposition that has been proved either by the user or elsewhere and that is stated for the expressed purpose of immediate use in the proof of another proposition.

Thus, in the development of the theory of probability and mathematical statistics, axioms are the fundamental truths that are to be accepted at face value and not proven. Theorems and their corollaries are statements deducible from the fundamental truths and other proven statements and thus are *derived truths*. Lemmas represent results that facilitate proofs of other statements that are of more fundamental interest.

We elaborate on the concept of a lemma, since our discussions will implicitly rely on lemmas more than any other type of statement, but we will generally choose not to exhaustively catalogue lemmas in the discussions. What constitutes a lemma and what does not depends on the problem context or on one's point of view. A fundamental integration result from calculus could technically be referred to as a lemma when used in a proof of a statement in mathematical statistics, while in the study of calculus, it might be referred to as a theorem to be proved in and of itself. Since our study will require numerous auxiliary results from algebra, calculus, and matrix theory, exhaustively cataloguing these results as lemmas would be cumbersome and, more importantly, not necessary given the prerequisites assumed for this course of study, namely, a familiarity with the basic concepts of algebra, univariate and multivariate calculus, and an introduction to matrix theory. We will have occasion to state a number of lemmas, but we will generally reserve this label for more exotic mathematical results that fall outside the realm of mathematics encompassed by the prerequisites.

A.3 Elements of Set Theory

In the study of probability and mathematical statistics, sets are the fundamental objects to which probability will be assigned, and it is important that the concept of a set, and operations on sets, be well understood. In this section we review some basic properties of and operations on sets. This begs the following question: What is meant by the term *set*? In modern axiomatic developments of set theory, the concept of a set is taken to be primitive and incapable of being defined in terms of more basic ideas. For our purposes, a more intuitive notion of a set will suffice, and we avoid the complexity of an axiomatic development of the theory (see Marsden,¹ Appendix A, for a brief introduction to the axiomatic development). We base our definition of a set on the intuitive definition originally proposed by the founder of set theory, Georg Cantor (1845–1918).²

Definition A.1 Set

A **set** is a collection of objects with the following characteristics:

1. All objects in the collection are *clearly defined*, so that it is evident which objects are members of the collection and which are not;
2. All objects are *distinguishable*, so that objects in the collection do not appear more than once;
3. *Order is irrelevant* regarding the listing of objects in the collection, so two collections that contain the same objects but are listed in different order are nonetheless the *same set*; and
4. *Objects in the collection can be sets themselves*, so that a *set of sets* can be defined.

The objects in the collection of objects comprising a set are its **elements**. The term **members** is also used to refer to the objects in the collection. In order to signify that an object belongs to a given set, the symbol \in , will be used in an expression such as $x \in A$, which is to be read “ x is an element (or member) of the set A .” If an object is *not* a member of a given set, then a slash will be used as $x \notin A$ to denote that “ x is not an element (or member) of the set A .” Note that the slash symbol, $/$, is used to indicate negation of a relationship. The characteristics of sets presented in Def. A.1 will be clarified and elaborated upon in examples and discussions provided in subsequent subsections.

¹J.E. Marsden, (1974), *Elementary Classical Analysis*, San Francisco: Freeman and Co.

²Cantor's original text reads: "Unter einer 'Menge' verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die 'Elemente' von M genannt werden) zu einem ganzen" (*Collected Papers*, p. 282). Our translation of Cantor's definition is, "By set we mean any collection, M , of clearly defined, distinguishable objects, m (which will be called elements of M), which from our perspective or through our reasoning we understand to be a whole."

Set-Defining Methods

Three basic methods are used in defining the objects in the collection constituting a given set: (1) **exhaustive listing**; (2) **verbal rule**; and (3) **mathematical rule**. An *exhaustive listing* requires that each and every object in a collection be individually identified either numerically, if the set is a collection of numbers, or by an explicit verbal description, if the collection is not of numbers. The object descriptions are conventionally separated by commas, and the entire group of descriptions is enclosed in brackets. The following are examples of sets defined by an exhaustive listing of the objects that are elements of the set:

Example A.1 $S_1 = \{\text{HEAD}, \text{TAIL}\}$

Here S_1 is the set of possible occurrences when tossing a coin into the air and observing its resting position. Note the set can be equivalently represented as $S_1 = \{\text{TAIL}, \text{HEAD}\}$. \square

Example A.2 $S_2 = \{1, 2, 3, 4, 5, 6\}$

Here S_2 is the set of positive integers from 1 to 6. Note that the set S_2 can be equivalently represented by listing the positive integers 1 to 6 in any order. \square

A *verbal rule* is a verbal statement of characteristics that only the objects that are elements of a given set possess and that can be used as a test to determine set membership. The general form of the verbal rule is $\{x : \text{verbal statement}\}$, which is to be read “the collection of all x for which *verbal statement* is true.” The following are examples of sets described by verbal rules:

Example A.3 $S_3 = \{x : x \text{ is a college student}\}$

Here S_3 is the set of college students. An individual is an element of the set S_3 iff (if and only if) he or she is a college student. \square

Example A.4 $S_4 = \{x : x \text{ is a positive integer}\}$

Here S_4 is the set of positive integers 1, 2, 3, ... A number is an element of the set S_4 iff it is a positive integer. \square

A *mathematical rule* is of the same general form as a verbal rule, except the verbal statement is replaced by a mathematical expression of some type. The general form of the mathematical rule is $\{x : \text{mathematical expression}\}$, which is to be read “the collection of all x for which *mathematical expression* is true.” The following are examples of sets described by mathematical rules:

Example A.5 $S_5 = \{x : x = 2k + 1, k = 0, 1, 2, 3, \dots\}$

Here S_5 is the set of odd positive integers. A number is an element of the set S_5 iff the number is equal to $2k + 1$ for some choice of $k = 0, 1, 2, 3, \dots$. \square

Example A.6 $S_6 = \{x : 0 \leq x \leq 1\}$

Here S_6 is the set of numbers greater than or equal to 0 but less than or equal to 1. A number is an element of the set S_6 iff it is neither less than 0 nor greater than 1. \square

The choice of method for describing the objects that constitute elements of a set depends on what is convenient and/or feasible for the case at hand. For example, exhaustive listing of the elements in set S_6 is impossible. On the other hand, there is some discretion that can be exercised, since, for example, a verbal rule could have adequately described the set S_5 , say as $S_5 = \{x : x \text{ is an odd positive integer}\}$. A mixing of the basic methods might also be used, such as $S_5 = \{x : x = 2k+1, k \text{ is zero or a positive integer}\}$. One can choose whatever method appears most useful in a given problem context.

Note that although our preceding examples of verbal and mathematical rules treat x as inherently one-dimensional, a vector interpretation of x is clearly permissible. For example, we can represent the set of points on the boundary or interior of a circle centered at $(0,0)$ and having radius 1 as

$$S_7 = \{(x_1, x_2) : x_1^2 + x_2^2 \leq 1\},$$

or we can represent the set of input-output combinations associated with a two-input Cobb-Douglas production function as

$$S_8 = \{(y, x_1, x_2) : y = b_0 x_1^{b_1} x_2^{b_2}, x_1 \geq 0, x_2 \geq 0\}$$

for given numerical values of b_0 , b_1 , and b_2 . Of course, the entries in the x -vector need not be numbers, as in the set

$$S_9 = \{(x_1, x_2) : x_1 \text{ is an economist, } x_2 \text{ is an accountant}\}.$$

Set Classifications

Sets are classified according to the number of elements they contain and whether the elements are countable. We differentiate between sets that have a finite number of elements and sets whose elements are infinite in number, referring to a set of the former type as a **finite set** and a set of the latter type as an **infinite set**. In terms of countability, sets are classified as being either countable or uncountable. Note that when we count objects, we intuitively place the objects in a one-to-one correspondence with the positive integers, i.e., we identify objects one by one and count "1, 2, 3, 4, ..." Thus, a **countable set** is one whose elements can be placed in a one-to-one correspondence with some or all of the positive integers—any other set is referred to as an **uncountable set**.

A finite set is, of course, always countable, and thus it would be redundant to use the term "countable finite set." Sets are thus either **finite**, **countably infinite**, or **uncountably infinite**. Of the sets S_1 through S_9 described earlier, S_1 , S_2 , S_3 , and S_9 are finite, S_4 and S_5 are countably infinite, and S_6 , S_7 , and S_8 are uncountably infinite (why?).

Special Sets, Set Operations, and Set Relationships

We now proceed to a number of definitions and illustrations establishing relationships between sets, mathematical operations on sets, and the notions of the universal and empty sets.

Definition A.2
Subset

A is a subset of *B*, denoted as $A \subset B$ and read *A* is contained in *B*, iff every element of *A* is also an element of *B*.

Definition A.3
Equality of sets

Set *A* is equal to set *B*, denoted as $A = B$, iff every element of *A* is also an element of *B*, and every element of *B* is also an element of *A*, i.e., iff $A \subset B$ and $B \subset A$.

Definition A.4
Universal set

The set containing all objects under consideration in a given problem setting, and from which all subsets are extracted, is the **universal set**.

Definition A.5
Empty or null set

The set containing no elements, denoted by \emptyset , is called the **empty**, or **null set**.

Definition A.6
Set difference

Given any two subsets *A* and *B* of a universal set, the set of all elements in *A* that are *not* in *B* is called the **set difference** between *A* and *B* and is denoted by $A - B$. If $A \subset B$, then $A - B = \emptyset$

Definition A.7
Complement

Let *A* be a subset of a universal set, Ω . The **complement** of the set *A* is the set of all elements in Ω that are not in *A*, and is denoted by \bar{A} . Equivalently, $\bar{A} = \Omega - A$.

Definition A.8
Union

Let *A* and *B* be any two subsets of a universal set, Ω . Then the **union** of the sets *A* and *B* is the set of all elements in Ω that are in *at least one* of the sets *A* or *B*; it is denoted by $A \cup B$.

Definition A.9
Intersection

Let *A* and *B* be any two subsets of a specified universal set, Ω . Then the **intersection** of the sets *A* and *B* is the set of all elements in Ω that are in *both* sets *A* and *B*, it is denoted by $A \cap B$.

Definition A.10
Mutually exclusive (or disjoint) sets

Subsets *A* and *B* of a universal set, Ω , are said to be **mutually exclusive** or **disjoint** sets iff they have no elements in common, i.e., iff $A \cap B = \emptyset$.

We continue to use the slash, $/$, to indicate negation of a relationship (recall that $/$ was previously used to indicate the negation of \in). Thus, $A \not\subset B$ denotes that A is not a subset of B , and $A \neq B$ denotes that A is not equal to B . We note here (and we shall state later as a theorem) that it is a logical requirement that \emptyset is a subset of any set A , since if \emptyset does not contain any elements, it cannot be the case that $\emptyset \not\subset A$, since the negation of \subset would require the existence of an element in \emptyset that was not in A .

Example A.7 Let the universal set be defined as $\Omega = \{x : 0 \leq x \leq 1\}$, and define three additional sets as

$$A = \{x : 0 \leq x \leq .5\}, \quad B = \{x : .25 \leq x \leq .75\} \quad \text{and,} \quad C = \{x : .75 < x \leq 1\}.$$

Then we can establish the following set relationships:

$$\bar{B} = \{x : 0 \leq x < .25 \text{ or } .75 < x \leq 1\},$$

$$A \cup C = \{x : 0 \leq x \leq .5 \text{ or } .75 < x \leq 1\},$$

$$\bar{B} \subset A \cup C,$$

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

$$\bar{C} = A \cup B = \{x : 0 \leq x \leq .75\},$$

$$A - B = \{x : 0 \leq x < .25\},$$

$$A \cap B = \{x : .25 \leq x \leq .5\}. \quad \square$$

Note that although our definitions of subset, equality of sets, set difference, complement, union, and intersection explicitly involve only two sets, A and B , it is implicit that the concepts can be applied to more complicated expressions involving an arbitrary number of sets. For example, since $A \cap B$ is itself a set, we can form its intersection with a set C as $(A \cap B) \cap C$, or form the set difference, $(A \cap B) - C$, or establish that $(A \cap B) =$ or $\neq C$, and so on. The point is that the concepts apply to sets, which themselves may have been constructed from other sets via various set operations.

Example A.8 Let Ω , A , B , and C be defined as in Ex. A.7. Then the following set relationships can be established:

$$A \cup B \cup C = \Omega = \{x : 0 \leq x \leq 1\},$$

$$(A \cup C) \cap B = \{x : .25 \leq x \leq .5\},$$

$$(A \cap B) \cap C = \overline{(A \cup B)} \cap \bar{C} = \emptyset,$$

$$(B \cup C) - A = \{x : .5 < x \leq 1\},$$

$$((A \cup B) - (A \cap B)) \subset \bar{C} \cap (\bar{A} \cup \bar{B}).$$

Can \subset be replaced with $=$ in the last relationship? \square

It is sometimes useful to conceptualize set relationships through illustrations called *Venn diagrams* (named after the nineteenth-century English logician, John Venn). In a Venn diagram, the universal set is generally denoted by a rectangle, with subsets of the universal set represented by various geometric shapes located within the bounds of the rectangle. Figure A.1 uses Venn diagrams to illustrate the set relationships defined previously.

Rules Governing Set Operations

Operations on sets must satisfy a number of basic rules. We state these basic rules as theorems, although we will not take the time to prove them here. The reader may wish to verify the plausibility of some of the theorems through the use of Venn diagrams. One of DeMorgan's laws will be proved to illustrate the formal proof method for the interested reader.

Theorem A.1 (Idempotency Laws)

$$A \cup A = A \quad \text{and} \quad A \cap A = A$$

Theorem A.2 (Commutative Laws)

$$A \cup B = B \cup A \quad \text{and} \quad A \cap B = B \cap A$$

Theorem A.3 (Associative Laws)

$$(A \cup B) \cup C = A \cup (B \cup C) \quad \text{and} \quad (A \cap B) \cap C = A \cap (B \cap C)$$

Theorem A.4 (Distributive Laws)

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad \text{and} \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Theorem A.5 (Identity Elements \cap and \cup)

$$A \cap \Omega = A \quad (\Omega \text{ is the identity element for } \cap)$$

$$A \cup \emptyset = A \quad (\emptyset \text{ is the identity element for } \cup)$$

Theorem A.6 (Intersection and Union of Complements)

$$A \cup \bar{A} = \Omega \quad \text{and} \quad A \cap \bar{A} = \emptyset$$

Theorem A.7 (Complements of Complements)

$$(\bar{\bar{A}}) = A$$

Theorem A.8 (Intersection with the Null Set)

$$A \cap \emptyset = \emptyset$$

Theorem A.9 (Null Set as a Subset) *If A is any set, then $\emptyset \subset A$*

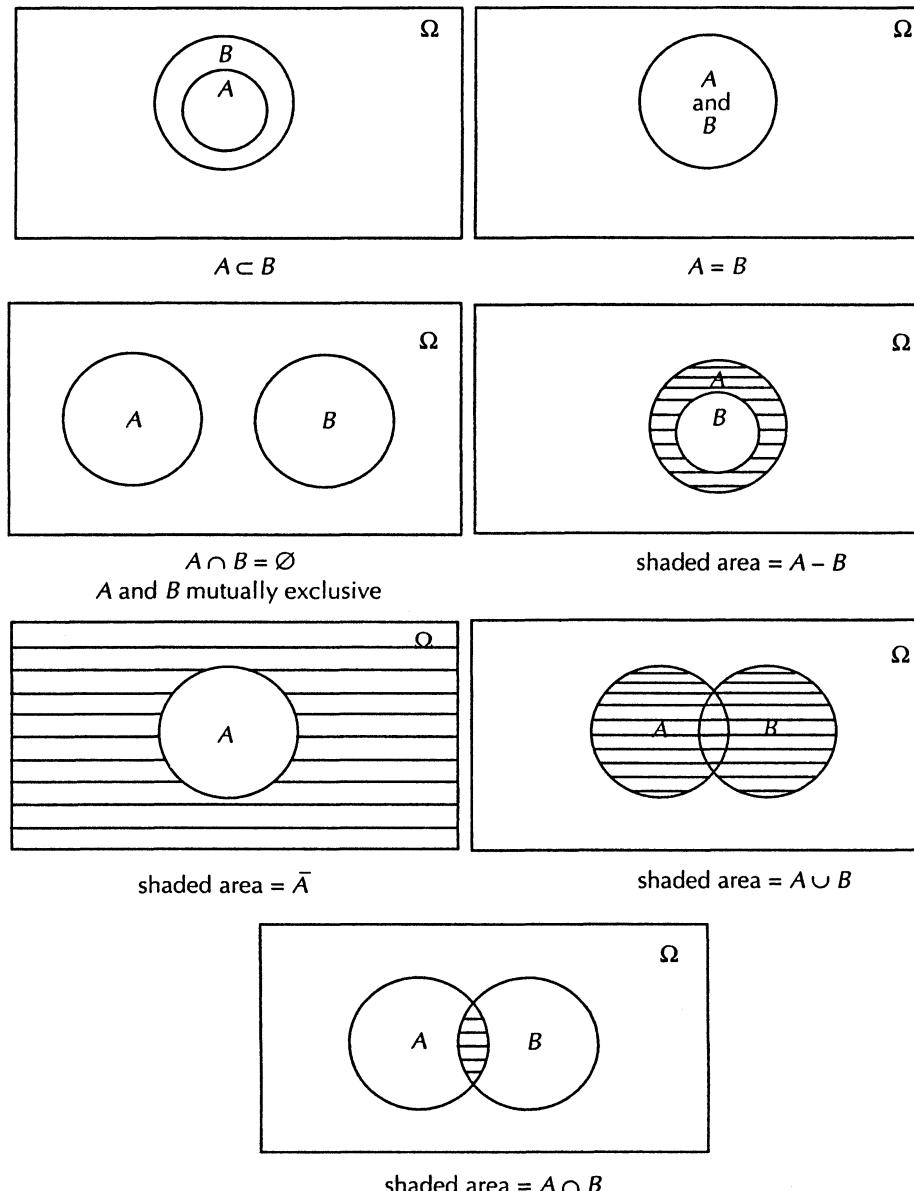


Figure A-1
Venn diagrams illustrating set relationships.

Theorem A.10 (DeMorgan's Laws)

$$\overline{A \cup B} = \bar{A} \cap \bar{B} \quad \text{and} \quad \overline{A \cap B} = \bar{A} \cup \bar{B}$$

Example A.9 Formal Proof of $\overline{A \cap B} = \bar{A} \cup \bar{B}$

By definition of the equality of sets, two sets are equal iff each is contained in the other. We first demonstrate that $\overline{A \cap B} \subset \bar{A} \cup \bar{B}$. By definition, $x \in \overline{A \cap B}$

implies that $x \notin A \cap B$. Suppose $x \notin \bar{A} \cup \bar{B}$. This implies $x \notin \bar{A}$ and $x \notin \bar{B}$, which implies $x \in A$ and $x \in B$, i.e., $x \in A \cap B$, a contradiction. Therefore, if $x \in \bar{A} \cap \bar{B}$, then $x \in \bar{A} \cup \bar{B}$, which implies $\bar{A} \cap \bar{B} \subset \bar{A} \cup \bar{B}$. We next demonstrate that $\bar{A} \cup \bar{B} \subset \bar{A} \cap \bar{B}$. Let $x \in \bar{A} \cup \bar{B}$. Then $x \notin A \cap B$, for if it were, then $x \in A$ and $x \in B$, contradicting that x belongs to at least one of \bar{A} and \bar{B} . However, $x \notin A \cap B$ implies $x \in \bar{A} \cap \bar{B}$, and thus $\bar{A} \cup \bar{B} \subset \bar{A} \cap \bar{B}$. \square

We remind the reader that since the sets used in Theorems A.1–A.10 could themselves be the result of set operations applied to other sets, the theorems are extendable in a myriad of ways to involve an arbitrary number of sets. For example, in the first of DeMorgan's laws listed in Theorem A.10, if $A = C \cup D$ and $B = E \cup F$, then by substitution,

$$\overline{(C \cup D \cup E \cup F)} = \overline{(C \cup D)} \cap \overline{(E \cup F)}.$$

Then by applying Theorem A.10 to both $(C \cup D)$ and $(E \cup F)$, we obtain a generalization of DeMorgan's law as

$$\overline{(C \cup D \cup E \cup F)} = \bar{C} \cap \bar{D} \cap \bar{E} \cap \bar{F}.$$

Given the wide range of extensions that are possible, Theorems A.1–A.10 provide a surprisingly broad conceptual foundation for applying the rules governing set operations.

Some Useful Set Notation Situations sometimes arise in which one is required to denote the union or intersection of a large number of sets. A convenient notation that represents such unions or intersections quite efficiently is available. Two types of notations are generally used, and they are differentiated on the basis of whether the union or intersection is of sets identified by a natural sequence of integer subscripts or whether the sets are identified by subscripts, say i 's, that are elements of some set of subscripts, I , called an **index set**.

Definition A.11

Multiple union notation

a. $\bigcup_{i=1}^n A_i = A_1 \cup A_2 \cup A_3 \dots \cup A_n$.

b. $\bigcup_{i \in I} A_i =$ union of all sets A_i for which $i \in I$.

Definition A.12

Multiple intersection notation

a. $\bigcap_{i=1}^n A_i = A_1 \cap A_2 \cap A_3 \dots \cap A_n$.

b. $\bigcap_{i \in I} A_i =$ intersection of all sets A_i for which $i \in I$.

Example A.10

Let the universal set be defined as $\Omega = \{x: 0 \leq x \leq 1\}$, and examine the following subsets of Ω :

$$A_1 = \{x: 0 \leq x \leq .25\}, \quad A_2 = \{x: 0 \leq x \leq .5\},$$

$$A_3 = \{x: 0 \leq x \leq .75\}, \quad A_4 = \{x: .75 \leq x \leq 1\}.$$

Define the index sets I_1 and I_2 as

$$I_1 = \{1, 3\} \text{ and } I_2 = \{1, 3, 4\}.$$

Then,

$$\cup_{i=1}^4 A_i = \cup_{i=2}^4 A_i = \cup_{i=3}^4 A_i = \{x : 0 \leq x \leq 1\} = \Omega,$$

$$\cup_{i \in I_1} A_i = A_1 \cup A_3 = \{x : 0 \leq x \leq .75\},$$

$$\cup_{i \in I_2} A_i = A_1 \cup A_3 \cup A_4 = \{x : 0 \leq x \leq 1\} = \Omega,$$

$$\cap_{i=1}^4 A_i = \cap_{i=2}^4 A_i = \emptyset,$$

$$\cap_{i=3}^4 A_i = \{.75\},$$

$$\cap_{i \in I_1} A_i = A_1 \cap A_3 = \{x : 0 \leq x \leq .25\},$$

$$\cap_{i \in I_2} A_i = A_1 \cap A_3 \cap A_4 = \emptyset.$$

□

Whenever a set A is an interval subset of the *real line* (where the *real line* refers to all of the numbers between $-\infty$ and ∞), the set can be indicated in abbreviated form by the standard notation for intervals, stated in the following definition.

Definition A.13
Interval set notation

Let a and b be two numbers on the real line for which $a < b$. Then the following four sets, called intervals with endpoints a and b , can be defined as

a. Closed interval:

$$[a, b] = \{x : a \leq x \leq b\},$$

b. Half-open (or half-closed) intervals:

$$(a, b] = \{x : a < x \leq b\}, \text{ and}$$

$$[a, b) = \{x : a \leq x < b\},$$

c. Open interval:

$$(a, b) = \{x : a < x < b\}.$$

Note that *weak inequalities*, $x \leq$ or $\leq x$, are signified by brackets $]$ or $[$, respectively. *Strong inequalities*, $x <$ or $< x$, are signified by parentheses $)$ or $($, respectively. Note further that whether the interval set contains its endpoints determines whether the set is closed.

As we have already done, (x, y) will also be used to denote coordinates in the two-dimensional plane. The context of the discussion will make clear whether we are referring to an open interval (a, b) or a pair of coordinates (x, y) .

A.4 Relations, Point Functions, and Set Functions

The concepts of point function and set function are central to our discussion of probability and statistics. We will see that probabilities can be represented

by set functions and that in a large number of cases of practical interest, set functions can in turn be represented by a summation or integration operation applied to point functions. While readers may be somewhat familiar with point functions from introductory courses in algebra, the concept of a set function may not be familiar. We will review both function concepts within the broader context of the theory of relations. The relations context facilitates the presentation of a very general definition of "function" in which inputs into and outputs from the function may be objects of any kind, including, but not limited to, numbers. The relations context also facilitates a demonstration of the significant similarities between the concept of a set function and the more familiar point function concept.

Cartesian Product

The concept of a relation can be made clear once we define what is meant by the Cartesian product of two sets A and B , named after the French mathematician René Descartes (1596–1650).

Definition A.14
Cartesian product of A and B

Let A and B be two sets. Then the **Cartesian product** of A and B , denoted as $A \times B$, is the set of ordered pairs

$$A \times B = \{(x, y) : x \in A, y \in B\}.$$

In words, $A \times B$ is the set of all possible pairs (x, y) such that x is an element of the set A and y is an element of the set B . Note carefully that the pairs are *ordered* in the sense that the first object in the pair must come from set A and the second object from set B .

Example A.11

Let $A = \{x : 1 \leq x \leq 2\}$ and $B = \{y : 2 \leq y \leq 4\}$. Then $A \times B = \{(x, y) : 1 \leq x \leq 2 \text{ and } 2 \leq y \leq 4\}$ (see Figure A.2). \square

Example A.12

Let $A = \{x : x \text{ is a man}\}$ and $B = \{y : y \text{ is a woman}\}$. Then $A \times B = \{(x, y) : x \text{ is a man and } y \text{ is a woman}\}$, which is the set of all possible man–woman pairings. \square

We will have use for a more general notion of Cartesian product involving more than just two sets. The extension is given in the next definition.

Definition A.15
Cartesian product (general)

Let A_1, \dots, A_n be n sets. Then the **Cartesian product** of A_1, \dots, A_n is the set of ordered n -tuples

$$\times_{i=1}^n A_i = A_1 \times A_2 \times \dots \times A_n = \{(x_1, \dots, x_n) : x_i \in A_i, i = 1, \dots, n\}.$$

In words, $\times_{i=1}^n A_i$ is the set of all possible n -tuples (x_1, \dots, x_n) such that x_1 is an element of set A_1 , x_2 is an element of set A_2 , and so on. Note that should the

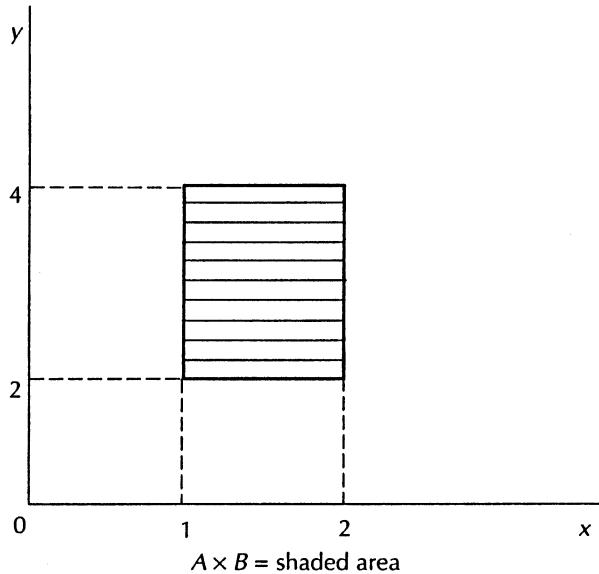


Figure A-2
 $A \times B = \text{shaded area.}$

need arise, a general Cartesian product of sets could also be represented by the notation $\times_{i \in I} A_i$, where here the product is taken over all sets having subscript i in the index set I (recall Def. A.11 and Def. A.12, and the use of index set notation).

In certain cases, we may be interested in forming a Cartesian product of a set A with itself. While we might represent such a Cartesian product by the notation

$$\times_{i=1}^n A = \{(x_1, \dots, x_n) : x_i \in A, i = 1, \dots, n\},$$

such a Cartesian product is generally denoted by A^n , and so, for example, $A^2 = A \times A$.

Relation (Binary)

We now define what we mean by the term *binary relation*.

Definition A.16
Binary Relation

Any subset of the Cartesian product $A \times B$ is a **binary relation from A to B** .

Note that the adjective *binary* signifies that only two sets are involved in the relation. Relations that involve more than two sets can be defined, but for our purposes the concept of a binary relation will suffice.³ Henceforth, we will use the word *relation* to mean binary relation. We should also mention that in the case where $B = A$, we will simply remain consistent with Def. A.16 and

³A higher-order relation could be defined by taking a subset of the Cartesian product $\times_{i=1}^n A_i$, for example.

refer to a subset of $A \times A$ as a *relation from A to A*, although in this special case some authors prefer to call the subset of $A \times A$ a *relation on A*.

Now let $S \subset A \times B$. Thus, by definition, S is a relation from A to B . (We emphasize at this point that the choice of the letter S is quite arbitrary, and we could just as well have chosen any other letter to represent a subset of $A \times B$ defining a relation from A to B .) If $(x, y) \in S$, we say that x is in the relation S to y or that x is S -related to y . An alternative notation for $(x, y) \in S$ is xSy . Also, we use $S: A \rightarrow B$ as an abbreviation for “the relation S from A to B .”

As it now stands, the concept of a relation no doubt appears quite abstract. However, in practice, it is the context provided by the definition of the subset S and the definitions of the sets A and B that provide intuitive meaning to xSy . That is, x will be S -related to y because of some property satisfied by the (x, y) pair, the property being indicated in the set definition of S . The real-world objects being related will be clearly identified in the set definitions of A and B . Some examples will clarify the intuitive side of the relation concept.

Example A.13 Let $A = [0, \infty)$, and form the Cartesian product $A^2 = \{(x, y) : x \in A \text{ and } y \in A\}$. The set A^2 can thus be interpreted as the nonnegative (or first) quadrant of the Euclidean plane. Then $S = \{(x, y) : x \geq y, (x, y) \in A^2\}$ is a relation from A to A representing the set of points in the nonnegative quadrant for which the first coordinate has a value greater than or equal to the value of the second coordinate. The defining property of the relation S is “ \geq .” This is displayed in Figure A.3. \square

Example A.14 Let $A = \{x : x \text{ is an employed U.S. citizen}\}$ and $B = \{y : y \text{ is a U.S. corporation}\}$. Then $A \times B = \{(x, y) : x \text{ is an employed U.S. citizen and } y \text{ is a U.S. corporation}\}$ is the set of all possible pairings of employed U.S. citizens with U.S. corporations.

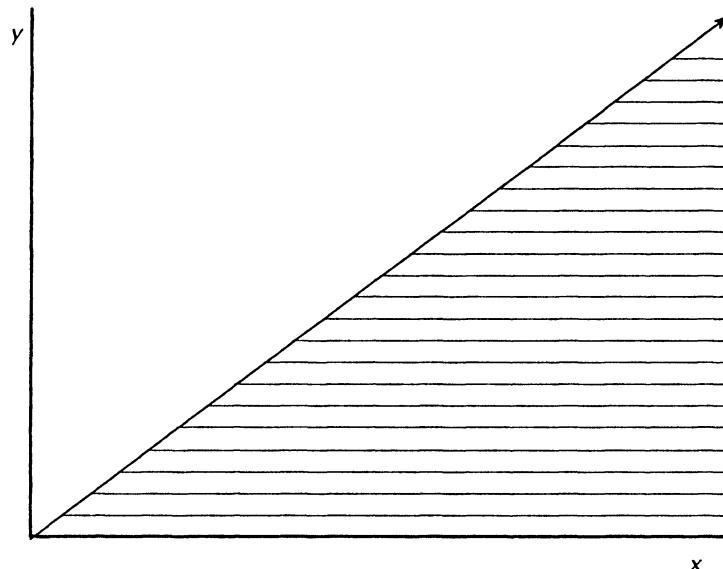


Figure A-3
 xSy in shaded area.

The relation $S = \{(x, y) : x \text{ is employed by } y, (x, y) \in A \times B\}$ from A to B is the collection of U.S. citizens who are employed by U.S. corporations paired with their respective corporate affiliation. The defining property of the relation is the phrase "is employed by," and xSy iff x is a U.S. citizen employed by a U.S. corporation, and y is his or her corporate affiliation. \square

Function

We are now in a position to define what is meant by the concept of a function. As indicated in the following definition, a function is simply a special type of relation. In the definition we introduce the symbol \forall , which means **for every** or **for all**, and the symbol \exists , which means **there exists**.

Definition A.17
Function

A function from A to B is a relation $S: A \rightarrow B$ such that $\forall a \in A \exists$ one unique $b \in B$ such that $(a, b) \in S$.

A relation satisfying the above condition will often be given a special symbol to distinguish the relation as a function. A popular symbol used for this purpose is " f ," where $f: A \rightarrow B$ is a common notation for designating the **function f from A to B** . As we had remarked when choosing a symbol to represent a relation, the choice of the letter f is arbitrary, and when it is convenient or useful, any other letter or symbol could be used to depict a subset of $A \times B$ that represents a function from A to B . In the text we will often have occasion to use a variety of letters to designate various functions of interest.

The unique element $b \in B$ that the function $f: A \rightarrow B$ associates with a given element $x \in A$ is called the **image of x under f** and is represented symbolically by the notation $f(x)$. If $f(x)$ is a real number, the image of x under f is alternatively referred to as the **value of the function f at x** . In the following example, we use R to denote the set of real numbers $(-\infty, \infty)$, i.e., R stands for the **real line**. Furthermore, the nonnegative subset of the real line is represented by $R_{\geq 0} = [0, \infty)$.

Example A.15 Let $f: R \rightarrow R_{\geq 0}$ be defined by $f = \{(x, y) : y = x^2, x \in R\}$. The **image of -2 under f** is $f(-2) = 4$. The **value of the function f at 3** is $f(3) = 9$. \square

Associated with a given function, f , are two important sets called the **domain** and **range** of the function.

Definition A.18
Domain and range of a function

The **domain** of a function $f: A \rightarrow B$ is defined as $D(f) = A$. The **range** of f is defined as $R(f) = \{y : y = f(x), x \in A\}$.

Thus, the domain of a function $f: A \rightarrow B$ is simply the set A of the Cartesian product $A \times B$ associated with the function. The range of f is the collection of all elements in B that are images of the elements in A under the function f . It

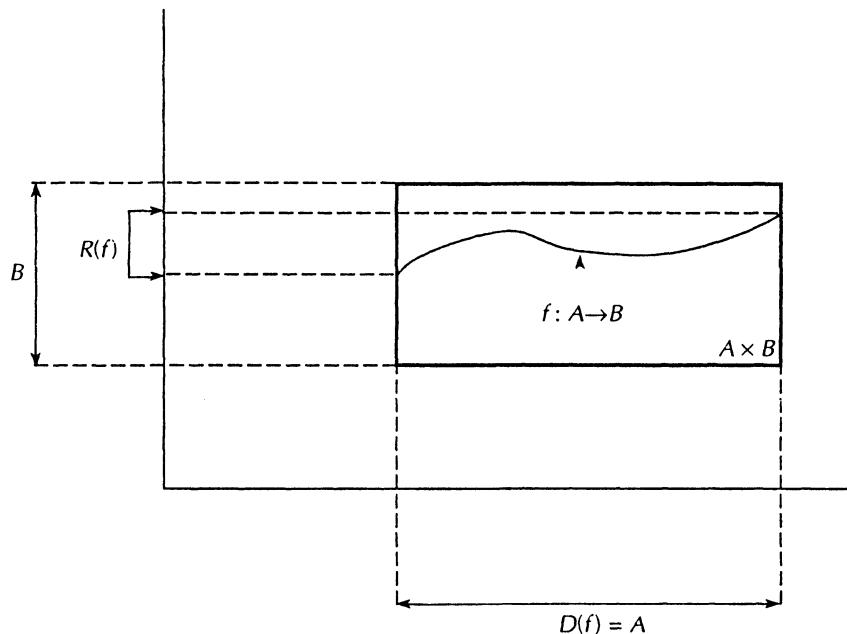


Figure A-4
Function with domain
 $D(f)$ and range $R(f)$

follows that $R(f) \subset B$. Figure A.4 provides a pictorial example of a function, including its *domain* and *range*.

Note that the concept of a function is completely general regarding the nature of the elements of the sets $D(f) = A$, $R(f)$, and B . The elements can be numbers, or other objects, or the elements can be sets themselves. For our work, it will suffice to deal only with *real-valued functions*, meaning that $R(f)$ is a set of real numbers.

Definition A.19
Real-valued function

A function $f: A \rightarrow B$ such that $R(f) \subset R$ is called a **real-valued function**.

The function defined in Ex. A.15 is a real-valued function. The following is another example.

Example A.16

Examine the Cobb-Douglas production function $f: R_{\geq 0}^2 \rightarrow R_{\geq 0}$ defined as $f = \{(x_1, x_2), y\}: y = 10x_1^2 x_2, (x_1, x_2) \in R_{\geq 0}^2\}$. Interpreting (x_1, x_2) as inputs into a production process, and y as the output of the process, we see that the domain of the production function is $D(f) = R_{\geq 0}^2$, i.e., any nonnegative level of the input pair (x_1, x_2) is an admissible input level. The associated range of the production function is $R(f) = R_{\geq 0}$, i.e., any nonnegative level of output, y , is possible. Since $R(f) = [0, \infty) \subset R$, the production function is a real-valued function. \square

In some cases the relation from A to B that defines the function $f: A \rightarrow B$ also defines a function from B to A . This relates to the concept of an inverse

function. In particular, if for each element $y \in B$ there exists precisely one element $x \in A$ whose image under f is y , then such an *inverse function* exists.

Definition A.20
Inverse function of f

Let $f: A \rightarrow B$ be a function from A to B . If $R(f) = B$ and $\forall y \in B \exists$ a unique $x \in A$ such that $y = f(x)$, then the relation $\{(y, x): y = f(x), y \in B\}$ is a function from B to A called the **inverse function of f** and denoted by $f^{-1}: B \rightarrow A$.

Note that neither of the functions in Ex. A.15 or Ex. A.16 is such that an inverse function exists. In Ex. A.15, the uniqueness condition of Def. A.20 is violated since $\forall y \neq 0$ there exist two values of x for which $y = x^2$, namely $x = \pm\sqrt{y}$. For example, when $y = 4$, $x = 2$ and -2 are each such that $y = x^2$. The reader can verify that Ex. A.16 also violates the uniqueness condition where an infinite number of (x_1, x_2) values satisfy $y = 10x_1^2x_2$ for a fixed value of y (defining level sets or isoquants of the production function). Also, note that an inverse function does not exist for the function illustrated in Figure A.4.

As an example of a function for which an inverse function does exist, consider the following.

Example A.17 Let $f: R \rightarrow R_+$ be defined as

$$f = \{(x, y): y = e^x, x \in R\}, \quad \text{where } R_+ = (0, \infty).$$

Note that the *inverse function* can be represented as

$$f^{-1} = \{(y, x): x = \ln y, y \in R_+\},$$

so that $f^{-1}: R_+ \rightarrow R$. It is clear that $\forall y \in R_+, \exists$ one and only one $x \in R$ such that $x = \ln y$. \square

The final concept concerning functions that we will review here is the **inverse image** of $y \in R(f)$ or of $H \subset R(f)$. The inverse image of y is the set of domain elements $x \in D(f)$ such that $y = f(x)$, i.e., the collection of all x values in the domain of f whose image under the function f is y . The inverse image of y can be represented as the set $\{x: f(x) = y\}$, and when the inverse function exists, the inverse image of y can be represented as the value $f^{-1}(y)$. In Ex. A.17, the inverse image of 5 is $f^{-1}(5) = \ln(5) = 1.6094$; in Ex. A.15, the inverse image of 4 is $\{-2, 2\}$; and in Ex. A.16, the inverse image of 3 is the isoquant $\{(x_1, x_2): 3 = 10x_1^2x_2, (x_1, x_2) \in R_{\geq 0}^2\}$. Similarly, the inverse image of $H \subset R(f)$ is the set of x values in the domain of f whose images under f equal some $y \in H$, i.e., the inverse image of H is $\{x: f(x) = y, y \in H\}$, and if the inverse function exists, $\{x: x = f^{-1}(y), y \in H\}$.

Real-Valued Point Versus Set Functions

Two types of functions—point functions and set functions—are utilized extensively in modern discussions of probability and mathematical statistics. The

reader should already have considerable experience with the application of real-valued point functions, since this type of function is the one that appears in elementary algebra and calculus courses and is central to discussions of utility, demand, production, and supply that the reader has encountered in his or her study of economic theory. Specifically, a **real-valued point function** is a real-valued function whose domain consists of a collection of points, where points are represented by coordinate vectors in R^n . We have encountered examples of this type of function previously in Ex. A.15 through Ex. A.17. A typical ordered pair associated with a real-valued point function is of the form (x, y) , where x is a vector in R^n and y is a real number in R .

A set function is more general than a point function in that its domain consists of a collection of sets rather than a collection of points.⁴ A typical ordered pair belonging to a **real-valued set function** would have the form (A, y) , where A is a set of some type of objects and y is a real number in R . If the sets in the domain of the set function are contained in R^n , i.e., they are sets of real numbers, then a real-valued set function assigns a real number to each set of points in its domain, whereas a real-valued point function assigns a real number to each point in its domain. A pictorial illustration of a set function contrasted with a point function is given in Figure A.5.

Examples of set functions are presented below.

Example A.18 Let $\Omega = \{1, 2, 3\}$, and let A be the collection of all of the subsets of Ω , i.e., $A = \{A_1, A_2, \dots, A_8\}$, where $A_1 = \{1\}$, $A_2 = \{2\}$, $A_3 = \{3\}$, $A_4 = \{1, 2\}$, $A_5 = \{1, 3\}$, $A_6 = \{2, 3\}$, $A_7 = \{1, 2, 3\}$, and $A_8 = \emptyset$.

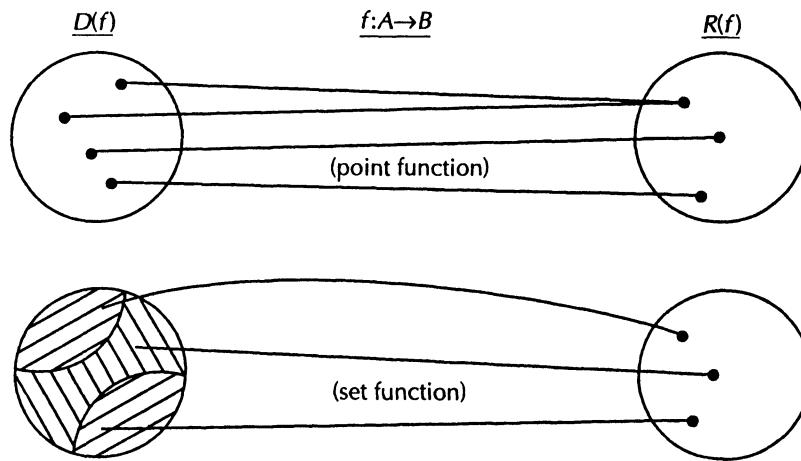


Figure A-5
Point versus set function.

⁴Note that, in a sense, a point function can be viewed as a special case of a set function, since points can be interpreted as singleton (single-element) sets. The set function concept is introduced to accommodate the case where one or more sets in its domain are not singleton.

The following is a real-valued set function $f: A \rightarrow R$:

$$f = \{(A_i, y) : y = \sum_{x \in A_i} x, A_i \subset A\},$$

where $\sum_{x \in A_i} x$ signifies the sum of the numerical values of all of the elements in the set A_i , and $\sum_{x \in \emptyset} x$ is defined to be zero. The range of the set function is $R(f) = \{0, 1, 2, 3, 4, 5, 6\}$, and the domain is the set of sets $D(f) = A$. The function can be represented in tabular form as follows:

| A_i | $f(A_i)$ |
|-------|----------|
| A_1 | 1 |
| A_2 | 2 |
| A_3 | 3 |
| A_4 | 3 |
| A_5 | 4 |
| A_6 | 5 |
| A_7 | 6 |
| A_8 | 0 |

□

Example A.19

Let $A = \{A_r : A_r = \{(x, y) : x^2 + y^2 \leq r^2\}, r \in [0, 1]\}$, so that A is a set of sets, where the typical element A_r represents the set of points in R^2 that are on the boundary and in the interior of a circle centered at $(0, 0)$ with radius r . The following is a real-valued set function $f: A \rightarrow R$:

$$f = \{(A_r, y) : y = \pi r^2, A_r \subset A\}.$$

Note that the set function assigns a real number representing the area to each set, A_r . The assignment is made for circles having a radius anywhere from 0 to 1. The range of the set function is given by $R(f) = [0, \pi]$, and the domain is the set of sets $D(f) = A$. □

A special type of set function called the size-of-set function will prove to be quite useful.

Definition A.21
Size of set function

Let A be any set of objects. The **size-of-set function**, N , is the set function that assigns to the set A the number of elements that are in set A , i.e., $N(A) = \sum_{x \in A} 1$.⁵

Applying the size-of-set function in Ex. A.18, note that $N(A) = 8$. In Ex. A.19, note that $N(A) = \infty$.

Another special (point) function that will be useful in our study is the **indicator function**, defined as follows.

⁵Note that $\sum_{x \in A} 1$ signified that a collection of 1's are being summed together, the number in the collection being equal to the number of elements in the set A . If $A = \emptyset$, effectively no 1's are being added together, and thus $N(\emptyset) = 0$.

Definition A.22
Indicator function

Let A be any subset of some universal set Ω . The indicator function, denoted by I_A , is a real-valued function with domain Ω and range $\{0, 1\}$ such that

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

Note that the indicator function *indicates* the set A by assigning the number 1 to any x that is an element of A , while assigning zero to any x that is not an element of A . The main use of the indicator function is notational efficiency in defining functions, as the following example illustrates.

Example A.20 Let the function $f: R \rightarrow R$ be defined by

$$f(x) = \begin{cases} 0 & \text{for } x \in (-\infty, 0] \\ x & \text{for } x \in (0, 2] \\ 3 - x & \text{for } x \in (2, 3] \\ 0 & \text{for } x \in (3, \infty) \end{cases}.$$

Utilizing the indicator function, we can alternatively represent $f(x)$ as

$$f(x) = xI_{(0,2]}(x) + (3 - x)I_{(2,3]}(x).$$

□

As a final note on the use of functions, we (as do the vast majority of other authors) will generally use a shorthand method for defining functions by simply specifying the relationship between elements in the domain of a function and their respective images in the range of the function. For example, we would define the function in Ex. A.19 by $f(A_r) = \pi r^2$ for $A_r \subset A$, or define the function in Ex. A.15 by $f(x) = x^2$ for $x \in R$. In all cases, the reader should remember that a function is a set of ordered pairs $(x, f(x))$, or $(A, f(A))$. The reader will sometimes find in the literature phrases like the function $f(x)$ or the set function $f(A)$. Literally speaking, such phrases are inconsistent, because $f(x)$ and $f(A)$ are not functions but rather *images* of elements in the domain of the respective functions. In fact, the reader should *not* take such phrases literally, but rather interpret these phrases as shorthand for phrases such as *the function whose values are given by $f(x)$* or *the set function whose values are given by $f(A)$* .

A.5 Combinations and Permutations

In a number of situations involving probability assignments, it will be useful to have an efficient method for counting the number of *different* ways a group of r objects can be selected from a group of n distinct objects, $n \geq r$. Obviously, if two groups of r objects do not contain the same r objects, they must be considered *different*. But what if two groups of r objects do contain the same objects, except the objects in the group are arranged in different orders? Are the two groups to be considered *different*? If difference in order constitutes difference in groups, then we are dealing with the notion of **permutations**. On the other hand, if the

order of listing the objects in a group is not used as a basis for distinguishing between groups, then we are dealing with the notion of **combinations**.

In order to establish a formula for determining the number of permutations of n distinct objects taken r at a time, the following example is suggestive.

Example A.21

Examine the number of different ways a group of three letters can be selected from the letters a, b, c, d , where difference in order of listing is taken to mean difference in groups. Note that the first letter can be chosen in four different ways. After we have chosen one of the letters for the first selection, the second selection can be any of the remaining three letters. Finally, after we have chosen two letters in the first two selections, there are then two letters left to be potentially chosen for the third selection. Thus, there are $4 \cdot 3 \cdot 2 = 24$ different ways of selecting a group of 3 letters from the letters a, b, c, d if difference in the order of listing constitutes difference in groups. (The reader should attempt to list the 24 different groups.) \square

The logic of the preceding example can be applied to establish a general formula for determining the number of permutations of n distinct objects taken r at a time:

$$\{n\}_r = \frac{n!}{(n-r)!} = n(n-1)(n-2)\dots(n-r+1),$$

where ! denotes the **factorial operation**, i.e.,

$$n! = n(n-1)(n-2)(n-3)\dots1.^6$$

Thus, for example, $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$. In Ex. A.21, $n = 4$ and $r = 3$, so that $\{4\}_3 = 4!/1! = 24$.

In order to establish a formula for determining the number of combinations of n distinct objects taken r at a time, we return to Ex. A.21.

Example A.21

(continued)

Examine the number of different ways a group of three letters can be selected from the letters a, b, c, d , where difference in order of listing does not imply the groups are different. Recall that we discovered that there were 24 permutations of the 4 letters a, b, c, d selected 3 at a time. Now note that any 3 letters, say a, b, c , can be arranged in $\{3\}_3 = 6$ different orders, which represents "overcounting" from the combinations point of view. Reducing the number of permutations by the degree of "overcounting" results in the number of combinations, i.e., there are $24/6 = 4$ combinations of the 4 letters taken 3 at a time, namely $\{a, b, c\}, \{a, b, d\}, \{a, c, d\}$, and $\{b, c, d\}$. \square

In the preceding example, the number of permutations of $n (= 4)$ objects taken $r (= 3)$ at a time was reduced by a factor of $r! (= 3!)$, where the latter value

⁶By definition, we take $0! = 1$.

represents the number of possible permutations of r objects. This suggests the general formula for the number of combinations of n objects taken r at a time:

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

In Ex. A.21, we have

$$\binom{4}{3} = \frac{4!}{1!3!} = 4$$

as the appropriate number of combinations.

The concept of combinations is useful in determining the number of subsets that can be constructed from a finite set A . Note that in counting the number of subsets, changes in the order of listing set elements do not produce a different set, e.g., the sets $\{a, b, c\}$ and $\{c, a, b\}$ are the same set of letters (recall the definition of a set). Then the total number of subsets of a set A containing n elements is given by the number of different subsets defined by taking no elements (i.e., the null set) plus the number of different subsets defined by taking one element, plus the number of different subsets defined by taking two elements, ..., and finally, the number of different subsets defined by taking all n elements (i.e., the set A itself). Thus, the total number of different subsets of A can be written as

$$\sum_{r=0}^n \binom{n}{r} = \sum_{r=0}^n \frac{n!}{(n-r)!r!}.$$

This sum can be greatly simplified by recalling that

$$(x+y)^n = \sum_{r=0}^n \binom{n}{r} x^r y^{n-r}, \quad n = 1, 2, \dots,$$

which is the **binomial theorem**. Then letting $x = y = 1$, we have that

$$2^n = \sum_{r=0}^n \binom{n}{r},$$

so that 2^n is the number of different subsets contained in a set A that has n elements.

Example A.22 In Ex. A.18, recall that we identified a total of eight subsets of the set $\Omega = \{1, 2, 3\}$. This is the number of subsets we would expect from our discussion above, i.e., since $n = 3$, there are $2^3 = 8$ subsets of Ω . \square

It should be noted that $\binom{n}{r}$ is *defined* to be 0 whenever $n < r$, or whenever n and/or $r < 0$ or are not integer valued. The rationale for $\binom{n}{r} = 0$ in each of these cases is that there is no way to define subsets of size r from a collection of n objects for the designated values of n and r .

When n is large, the calculation of $n!$ needed in the previous formulas pertaining to numbers of permutations or combinations can be quite formidable.

A result known as Stirling's formula can provide a useful approximation to $n!$ for large n :

Definition A.23
*Stirling's formula*⁷

$$n! \approx (2\pi)^{1/2} n^{n+5} e^{-n} \text{ for large } n.$$

A logical question to ask regarding the use of Stirling's formula is how large is "large n "? Stirling's formula invariably underestimates $n!$, but the percentage error $\leq 1\%$ for $n \geq 10$, and it monotonically decreases as $n \rightarrow \infty$.

A.6 Summation, Integration and Matrix Differentiation Notation

We will use a number of variations on summation and integration notation in this text. The meaning of the various types of notation are presented in Table A.1.

We illustrate the use of some of the notation in the following example.

Example A.23

Summation Notation

Let $A_1 = \{1, 2, 3\}$, $A_2 = \{2, 4, 6\}$, $A = A_1 \times A_2$, $B = \{(x_1, x_2) : x_1 \in A_1, x_2 \in \{x_1, x_1 + 1, \dots, 3x_1\}\}$, $y = \{y_1, y_2, \dots, y_n\}$, and $f(x_1, x_2) = x_1 + 2x_2^2$. Then

$$\sum_{x \in A_1} x = 1 + 2 + 3 = 6, \quad \sum_{x \in A_2} x^2 = 2^2 + 4^2 + 6^2 = 56,$$

$$\sum_{i \in A_2} y_i = y_2 + y_4 + y_6, \quad \sum_{i \in A_1} y_i = \sum_{i=1}^3 y_i = y_1 + y_2 + y_3,$$

$$\sum_{x_1 \in A_1} \sum_{x_2 \in A_2} f(x_1, x_2) = \sum_{x_1 \in A_1} \sum_{x_2 \in A_2} (x_1 + 2x_2^2) = 354,$$

$$\sum_{(x_1, x_2) \in A} f(x_1, x_2) = \sum_{x_1 \in A_1} \sum_{x_2 \in A_2} f(x_1, x_2) = 354,$$

$$\sum_{(x_1, x_2) \in B} f(x_1, x_2) = \sum_{x_1=1}^3 \sum_{x_2=x_1}^{3x_1} (x_1 + 2x_2^2) = 802. \quad \square$$

Example A.24

Integration Notation

Let $A_1 = [0, 3]$, $A_2 = [2, 4]$, $A = A_1 \times A_2$, $B = \{(x_1, x_2) : x_1 \in A_1, 0 < x_2 < x_1^2\}$, and $f(x_1, x_2) = x_1 x_2^2$. Then

$$\int_{x \in A_1} 2x \, dx = \int_0^3 2x \, dx = \frac{2x^2}{2} \Big|_0^3 = 9,$$

⁷See W. Feller (1968), *An Introduction to the Theory of Probability and Its Applications*, 3rd ed. pp. 52–54.

⁸Note that \approx means "approximately equal to."

Table A.1 Summation and Integration Notation

| Notation | Definition |
|--|--|
| $\sum_{i=\ell}^n x_i$ | Sum the values of $x_\ell, x_{\ell+1}, \dots, x_n$, i.e., $x_\ell + x_{\ell+1} + \dots + x_n$. |
| $\sum_{i \in I} x_i$ | Sum the values of the x_i 's, for $i \in I$. |
| $\sum_{x \in A} x$ | Sum the values of $x \in A$. |
| $\sum_{x=a}^b x$ | Sum the values of x in the sequence of integers from a to b , i.e., $a + (a+1) + (a+2) + \dots + b$. |
| $\sum_{i=\ell}^n \sum_{j=k}^m x_{ij}$ | Sum the values of the x_{ij} 's for $i = \ell, \ell+1, \dots, n$ and $j = k, k+1, \dots, m$. |
| $\sum_{i \in I} \sum_{j \in J} x_{ij}$ or $\sum \sum_{(i,j) \in A} x_{ij}$ | Sum the values of the x_{ij} 's for $i \in I$ and $j \in J$, or for $(i,j) \in A$. |
| $\sum_{x_1 \in A_1} \dots \sum_{x_n \in A_n} f(x_1, \dots, x_n)$ | Sum the values of $f(x_1, \dots, x_n)$ for $x_i \in A_i, i = 1, \dots, n$. |
| $\sum \dots \sum_{(x_1, \dots, x_n) \in A} f(x_1, \dots, x_n)$ | Sum the values of $f(x_1, \dots, x_n)$ for $(x_1, \dots, x_n) \in A$. |
| $\sum_{x_1=a_1}^{b_1} \dots \sum_{x_n=a_n}^{b_n} f(x_1, \dots, x_n)$ | Sum the values of $f(x_1, \dots, x_n)$ for x_i in the sequence of integers a_i to $b_i, i = 1, \dots, n$. |
| $\int_a^b f(x) dx$ | Integral of the function $f(x)$ from a to b (a can be $-\infty$ and/or b can be ∞). |
| $\int_{x \in A} f(x) dx$ | Integral of the function $f(x)$ over the set of points A . |
| $\int_{x_1 \in A_1} \dots \int_{x_n \in A_n} f(x_1, \dots, x_n) dx_n \dots dx_1$ | Iterated integral of the function $f(x_1, \dots, x_n)$ over the points $x_i \in A_i, i = 1, \dots, n$. |
| $\int \dots \int_{(x_1, \dots, x_n) \in A} f(x_1, \dots, x_n) dx_1 \dots dx_n$ | Multiple integral of the function $f(x_1, \dots, x_n)$ over the points $(x_1, \dots, x_n) \in A$. |
| $\int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \dots dx_1$ | Iterated integral of the function $f(x_1, \dots, x_n)$ for x_i in the (open, half open-half closed, or closed) interval a_i to b_i , for $i = 1, \dots, n$. |

$$\int_{x \in A_1 \cap A_2} x^2 dx = \int_2^3 x^2 dx = \frac{x^3}{3} \Big|_2^3 = \frac{19}{3},$$

$$\begin{aligned} \int_{x_1 \in A_1} \int_{x_2 \in A_2} f(x_1, x_2) dx_2 dx_1 &= \int_0^3 \int_2^4 f(x_1, x_2) dx_2 dx_1 \\ &= \int_0^3 \frac{x_1 x_2^3}{3} \Big|_2^4 dx_1 = \int_0^3 \frac{56}{3} x_1 dx_1 \\ &= \frac{56 x_1^2}{6} \Big|_0^3 = 84, \end{aligned}$$

$$\iint_{(x_1, x_2) \in A} f(x_1, x_2) dx_1 dx_2 = \int_{x_1 \in A_1} \int_{x_2 \in A_2} f(x_1, x_2) dx_2 dx_1 = 84,$$

$$\begin{aligned}
 \iint_{\{(x_1, x_2) \in B\}} f(x_1, x_2) dx_1 dx_2 &= \int_0^3 \int_0^{x_1^2} f(x_1, x_2) dx_2 dx_1 \\
 &= \int_0^3 \frac{x_1 x_2^3}{3} \Big|_0^{x_1^2} dx_1 = \int_0^3 \frac{x_1^7}{3} dx_1 \\
 &= \frac{x_1^8}{24} \Big|_0^3 = 273.375. \quad \square
 \end{aligned}$$

Regarding matrix differentiation notation, we utilize the following conventions. Let $g(\mathbf{x})$ and $\mathbf{y}(\mathbf{x})$ be a scalar and $n \times 1$ vector function of the $k \times 1$ vector \mathbf{x} , respectively. Then

| Derivative | Matrix Dimension | (i, j) th Entry |
|---|------------------|--|
| $\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}$ | $k \times 1$ | $\frac{\partial g}{\partial x_i}$ |
| $\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}'} = \frac{\partial g(\mathbf{x})'}{\partial \mathbf{x}}$ | $1 \times k$ | $\frac{\partial g}{\partial x_j}$ |
| $\frac{\partial g(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'}$ | $k \times k$ | $\frac{\partial^2 g}{\partial x_i \partial x_j}$ |
| $\frac{\partial \mathbf{y}(\mathbf{x})}{\partial \mathbf{x}}$ | $k \times n$ | $\frac{\partial \mathbf{y}_i}{\partial x_i}$ |
| $\frac{\partial \mathbf{y}(\mathbf{x})}{\partial \mathbf{x}'} = \frac{\partial \mathbf{y}(\mathbf{x})'}{\partial \mathbf{x}}$ | $n \times k$ | $\frac{\partial \mathbf{y}_i}{\partial x_j}$ |

Key Words, Phrases, and Symbols

| | | |
|--------------------------|--------------------------------|---|
| definition | subset | complements of complements |
| axiom (or postulate) | equality of sets, $=$ | intersection with null set |
| theorem (or proposition) | universal set | null set as a subset |
| corollary | empty or null set, \emptyset | DeMorgan's laws |
| lemma | set difference, $-$ | index set |
| set | complement, \bar{A} | multiple union notation |
| element | union, \cup | multiple intersection notation |
| \in | intersection, \cap | real line |
| negation, / | mutually exclusive (disjoint) | interval set notation |
| exhaustive listing | Venn diagram | closed, open, half-open-half-closed intervals |
| verbal rule | idempotency laws | Cartesian product |
| mathematical rule | commutative laws | binary relation from A to B , |
| finite set | associative laws | $S: A \rightarrow B$ |
| iff (if and only if) | distributive laws | xSy |
| infinite set | identity elements | \forall (for every) |
| countable set | intersection and union of | \exists (there exists) |
| uncountable set | complements | |

| | | |
|---|------------------------------|------------------------------|
| function from A to B , $f: A \rightarrow B$ | point function | permutations, $\{n\}_r$ |
| real-valued function | set function | combinations, $\binom{n}{r}$ |
| $A^2 = A \times A$ | size-of-set function | binomial theorem |
| $R_{\geq 0}, R_+$ | indicator function, $I_A(x)$ | Stirling's formula |
| image of x under f | "the function $f(x)$ " | summation notation |
| inverse function, $f^{-1}: B \rightarrow A$ | "the set function $f(A)$ " | integration notation |
| inverse image | | |

Problems

1. Using either an exhaustive listing, verbal rule, or mathematical rule, define the following sets:

- a. the set of all senior citizens receiving social security payments in the United States;
- b. the set of all positive numbers that are positive integer powers of the number 10 (i.e., $10^1, 10^2$, etc.);
- c. the set of all possible outcomes resulting from rolling a red and a green die and calculating the values of $y - x$, where y = number of dots on the red die and x = number of dots on the green die;
- d. the set of all two-tuples (x_1, x_2) where x_1 is any real number and x_2 is related to x_1 by raising the number e to the power x_1 .

2. Label the sets you have identified in Problem 1. as being either finite, countably infinite, or uncountably infinite, and explain your choice.

3. For each set below, state whether the set is finite, countably infinite, or uncountably infinite.

- a. $S = \{x: x \text{ is a U.S. citizen who has purchased a Japanese car during the past year}\}.$
- b. $S = \{(x, y): y \leq x^2, x \text{ is a positive integer}, y \in R_{\geq 0}\}.$
- c. $S = \{p: p \text{ is the price of a quart of milk sold at a retail store in the United States on Friday, September 13, 1991}\}.$
- d. $S = \{x: x = 2y, y \text{ is a positive integer}\}.$

4. Let the universal set be $\Omega = [0, 10]$, and define the following subsets of Ω :

$$A = [0, 2], \quad B = [2, 7], \quad C = [5, 6], \quad D = \{2\}, \\ E = \{x: x = y^{-1}, y \text{ is an even positive integer } \geq 4\}.$$

- a. Define the following sets:

$$A \cup B, \quad A \cap B, \quad \overline{A \cup C}, \quad (A \cup D) \cap B, \\ B - C, \quad A \cap E, \quad \bar{D} \cap B.$$

b. For each of the sets in {a}, indicate whether the set is finite, countably infinite, or uncountably infinite.

5. Let the universal set be defined by $\Omega = [-5, 5]$, and define the following subsets of Ω :

$$A_1 = [-2, 1],$$

$$A_2 = \{1, 2\},$$

$$A_3 = [2, 5],$$

$$A_4 = [-5, -2].$$

Also, define an index set $I = \{1, 3, 4\}$.

- a. Define $\cup_{i \in I} A_i$.
- b. Define $\cup_{i=1}^4 A_i$.
- c. Define $A_1 \cap A_2$.
- d. Define $A_4 - A_1$.
- e. Define \bar{A}_4 .

6. Define the universal set, Ω , as

$$\Omega = \{x: 0 \leq x \leq 5 \text{ or } 10 \leq x \leq 20\},$$

and define the following subsets of Ω as

$$A_1 = \{x: 0 \leq x < 2.5\},$$

$$A_2 = \{x: 15 < x \leq 20\},$$

$$A_3 = \{x: 2.5 \leq x \leq 5 \text{ or } 10 \leq x \leq 20\},$$

$$A_4 = \{x: 0 \leq x \leq 5 \text{ or } 10 \leq x \leq 15\}.$$

In addition, define the following two index sets as

$$I_1 = \{1, 3\}, \quad I_2 = \{1, 4\}.$$

Define the following sets:

- a. $\cup_{i \in I_1} A_i$
- b. $\cap_{i=1}^4 A_i$
- c. $\cap_{i \in I_2} A_i$
- d. $\cap_{i=1}^2 A_i$

- e. $A_1 - A_2$
f. $A_4 - A_3$
g. \bar{A}_3
h. $A_2 - \bigcup_{i \in I} A_i$

7. In each situation below indicate whether the relation is a function. If so, determine the domain and range of the function.

- a. $A = [0, 10], B = [0, \ln(11)], S = \{(x, y) : y = \ln(1+x), (x, y) \in A \times B\}$.
c. $A = R_{\geq 0}^2, B = [0, \infty), S = \{(x_1, x_2, y) : y = 5x_1 x_2^2, (x_1, x_2, y) \in A \times B\}$.

8. For each relation below, state whether the relation is a function, and state whether an inverse function exists. Explicitly define the inverse function if it exists.

- a. Let $P = \{0.01, 0.02, \dots, 1.00, 1.01, \dots\}$ represent a set of possible prices for a given commodity, and let $Q = [0, \infty)$ represent possible levels of quantity demanded. Define $S: P \rightarrow Q$ as

$$S = \{(p, q) : q = 20p^{-1.5}, (p, q) \in P \times Q\}.$$

- b. Let $A = \{D : D = \{(x_1, x_2) : x_1 \in [a_1, b_1], x_2 \in [a_2, b_2]\}, a_1 < b_1, a_2 < b_2\}$ be a set of rectangular sets. Define $S: A \rightarrow R$ as

$$S = \{(D, y) : y = \text{area of } D, (D, y) \in A \times R_+\}.$$

9. Let $A = [0, \infty)$, and examine the following relation on A :

$$S = \{(x, y) : y = 2 + 3x, (x, y) \in A^2\}.$$

- a. Is S a function?
b. Is S a function?
c. Does an inverse function exist?
d. If you can, define $f(2)$ and $f^{-1}(5)$.
e. What is $D(f)$? What is $R(f)$?

10. Define a universal set as

$$\Omega = \{x : 0 \leq x \leq 5\},$$

and consider the set function

$$P(A) = .5 \int_{x \in A} x dx + 12.5,$$

where the domain of the set function is all subsets $A \subset \Omega$ of the form $A = [a, b]$, for $0 \leq a \leq b \leq 5$.

- a. What is the image of the set $A = [0, 2]$ under P ?
b. What is the image of Ω under P ?
c. What is the image of the set $A = [3, 3]$ under P ?
d. What is the range of the set function?
e. What is the inverse image of 2?

11. Define a set function that will assign the appropriate area to all rectangles of the form $[x_1, x_2] \times [y_1, y_2]$, $x_2 \geq x_1$ and $y_2 \geq y_1$, contained in R^2 . Be sure to identify the domain and range of the set function.

12. A statistics class has 20 students in attendance in a room where 25 desks are available for the students. How many different ways can the students leave 5 desks unoccupied?

13. There are 15 students in an econometrics class that you are attending.

- a. How many different ways can a three-person committee be formed to give a class report?
b. Of the number of possible three-person committees indicated in (a), how many involve you?

14. Competing for the title of Miss America are 50 contestants from each of the 50 states plus 1 contestant from the District of Columbia. How many different ways can the contestants be assigned the titles of Miss America, 1st runner up, ..., 4th runner up?

15. Let $A_1 = \{x : x \text{ is a positive integer}\}$, $A_2 = \{1, 2, 3, 4, 5\}$, $B = \{(x_1, x_2) : (x_1, x_2) \in A_1 \times A_2, x_1 \leq x_2\}$, and $y_i = i^2$. Calculate the values of the following sums.

- a. $\sum_{x \in A_2} x$
b. $\sum_{i \in A_2} y_i$
c. $\sum_{x_1 \in A_1} \sum_{x_2 \in A_2} (1/2)^{x_1} x_2^2$
d. $\sum_{x \in A_1 - A_2} (1/3)^x$
e. $\sum \sum_{(x_1, x_2) \in B} (x_1 + x_2)$

16. Let $A_1 = [0, \infty)$, $A_2 = [1, 10]$, and $B = \{(x_1, x_2) : (x_1, x_2) \in A_1 \times A_2, x_2 > x_1\}$. Calculate the values of the following integrals.

- a. $\int_{x \in A_1} (1/2)e^{-x/2} dx$
b. $\int_{x_1 \in A_1} \int_{x_2 \in A_2} x_2 e^{-x_1} dx_2 dx_1$
c. $\iint_{(x_1, x_2) \in B} (x_1 + x_2) dx_1 dx_2$
d. $\int_0^2 \int_{x_2 \in A_1 \cap A_2} x_1 x_2^2 dx_2 dx_1$

B

Useful Tables



- B.1 Cumulative Normal Distribution
- B.2 Student's t Distribution
- B.3 Chi-square Distribution
- B.4 F -Distribution: 5% Points
- B.5 F -Distribution: 1% Points



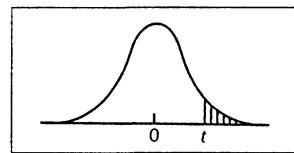
Table B.1 Cumulative Normal Distribution $F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$

| x | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| .0 | .5000 | .5040 | .5080 | .5120 | .5160 | .5199 | .5239 | .5279 | .5319 | .5359 |
| .1 | .5398 | .5438 | .5478 | .5517 | .5557 | .5596 | .5636 | .5675 | .5714 | .5753 |
| .2 | .5793 | .5832 | .5871 | .5910 | .5948 | .5987 | .6026 | .6064 | .6103 | .6141 |
| .3 | .6179 | .6217 | .6255 | .6293 | .6331 | .6368 | .6406 | .6443 | .6480 | .6517 |
| .4 | .6554 | .6591 | .6628 | .6664 | .6700 | .6736 | .6772 | .6808 | .6844 | .6879 |
| .5 | .6915 | .6950 | .6985 | .7019 | .7054 | .7088 | .7123 | .7157 | .7190 | .7224 |
| .6 | .7257 | .7291 | .7324 | .7357 | .7389 | .7422 | .7454 | .7486 | .7517 | .7549 |
| .7 | .7580 | .7611 | .7642 | .7673 | .7704 | .7734 | .7764 | .7794 | .7823 | .7852 |
| .8 | .7881 | .7910 | .7939 | .7967 | .7995 | .8023 | .8051 | .8078 | .8106 | .8133 |
| .9 | .8159 | .8186 | .8212 | .8238 | .8264 | .8289 | .8315 | .8340 | .8365 | .8389 |
| 1.0 | .8413 | .8438 | .8461 | .8485 | .8508 | .8531 | .8554 | .8577 | .8599 | .8621 |
| 1.1 | .8643 | .8665 | .8686 | .8708 | .8729 | .8749 | .8770 | .8790 | .8810 | .8830 |
| 1.2 | .8849 | .8869 | .8888 | .8907 | .8925 | .8944 | .8962 | .8980 | .8997 | .9015 |
| 1.3 | .9032 | .9049 | .9066 | .9082 | .9099 | .9115 | .9131 | .9147 | .9162 | .9177 |
| 1.4 | .9192 | .9207 | .9222 | .9236 | .9251 | .9265 | .9279 | .9292 | .9306 | .9319 |
| 1.5 | .9332 | .9345 | .9357 | .9370 | .9382 | .9394 | .9406 | .9418 | .9429 | .9441 |
| 1.6 | .9452 | .9463 | .9474 | .9484 | .9495 | .9505 | .9515 | .9525 | .9535 | .9545 |
| 1.7 | .9554 | .9564 | .9573 | .9582 | .9591 | .9599 | .9608 | .9616 | .9625 | .9633 |
| 1.8 | .9641 | .9649 | .9656 | .9664 | .9671 | .9678 | .9686 | .9693 | .9699 | .9706 |
| 1.9 | .9713 | .9719 | .9726 | .9732 | .9738 | .9744 | .9750 | .9756 | .9761 | .9767 |
| 2.0 | .9772 | .9778 | .9783 | .9788 | .9793 | .9798 | .9803 | .9808 | .9812 | .9817 |
| 2.1 | .9821 | .9826 | .9830 | .9830 | .9834 | .9838 | .9846 | .9850 | .9854 | .9857 |
| 2.2 | .9861 | .9864 | .9868 | .9871 | .9875 | .9878 | .9881 | .9884 | .9887 | .9890 |
| 2.3 | .9893 | .9896 | .9898 | .9901 | .9904 | .9906 | .9909 | .9911 | .9913 | .9916 |
| 2.4 | .9918 | .9920 | .9922 | .9925 | .9927 | .9929 | .9931 | .9932 | .9934 | .9936 |
| 2.5 | .9938 | .9940 | .9941 | .9943 | .9945 | .9946 | .9948 | .9949 | .9951 | .9952 |
| 2.6 | .9953 | .9955 | .9956 | .9957 | .9959 | .9960 | .9961 | .9962 | .9963 | .9964 |
| 2.7 | .9965 | .9966 | .9967 | .9968 | .9969 | .9970 | .9971 | .9972 | .9973 | .9974 |
| 2.8 | .9974 | .9975 | .9976 | .9977 | .9978 | .9979 | .9979 | .9979 | .9980 | .9981 |
| 2.9 | .9981 | .9982 | .9982 | .9983 | .9984 | .9984 | .9985 | .9985 | .9986 | .9986 |
| 3.0 | .9987 | .9987 | .9987 | .9988 | .9988 | .9989 | .9989 | .9989 | .9990 | .9990 |
| 3.1 | .9990 | .9991 | .9991 | .9991 | .9992 | .9992 | .9992 | .9992 | .9993 | .9993 |
| 3.2 | .9993 | .9993 | .9994 | .9994 | .9994 | .9994 | .9994 | .9995 | .9995 | .9995 |
| 3.3 | .9995 | .9995 | .9995 | .9996 | .9996 | .9996 | .9996 | .9996 | .9996 | .9997 |
| 3.4 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9998 |

(Source: Reprinted, by permission of the publisher, from A. M. Mood, F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, 3d ed., New York: McGraw-Hill, 1974, p. 552.)

Table B.2 Student's t Distribution

The first column lists the number of degrees of freedom (v). The headings of the other columns give probabilities (P) for t to exceed the entry value. Use symmetry for negative t values.



| $\frac{P}{v}$ | .10 | .05 | .025 | .01 | .005 |
|---------------|-------|-------|--------|--------|--------|
| 1 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 |
| 2 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 |
| 3 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 |
| 4 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 |
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 |
| 30 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 |
| 40 | 1.303 | 1.684 | 2.021 | 2.423 | 2.704 |
| 60 | 1.296 | 1.671 | 2.000 | 2.390 | 2.660 |
| 120 | 1.289 | 1.658 | 1.980 | 2.358 | 2.617 |
| ∞ | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 |

(Source: Reprinted, by permission of the publisher, from P. G. Hoel, *Introduction to Mathematical Statistics*, 4th ed., New York: John Wiley and Sons, Inc., 1971, p. 393.)

Table B.3 Chi-square Distribution

The first column lists the number of degrees of freedom (v). The headings of the other columns give probabilities (P) for the χ^2_v random variable to exceed the entry value.

| $v \setminus P$ | 0.995 | 0.990 | 0.975 | 0.950 | 0.900 | 0.750 |
|-----------------|--------------------------|-------------------------|-------------------------|-------------------------|-----------|-----------|
| 1 | 392704×10^{-10} | 157088×10^{-9} | 982069×10^{-9} | 393214×10^{-8} | 0.0157908 | 0.1015308 |
| 2 | 0.0100251 | 0.0201007 | 0.0506356 | 0.102587 | 0.210720 | 0.575364 |
| 3 | 0.0717212 | 0.114832 | 0.215795 | 0.351846 | 0.584375 | 1.212534 |
| 4 | 0.206990 | 0.297110 | 0.484419 | 0.710721 | 1.063623 | 1.92255 |
| 5 | 0.411740 | 0.554300 | 0.831211 | 1.145476 | 1.61031 | 2.67460 |
| 6 | 0.675727 | 0.872085 | 1.237347 | 1.63539 | 2.20413 | 3.45460 |
| 7 | 0.989265 | 1.239043 | 1.68987 | 2.16735 | 2.83311 | 4.25485 |
| 8 | 1.344419 | 1.646482 | 2.17973 | 2.73264 | 3.48954 | 5.07064 |
| 9 | 1.734926 | 2.087912 | 2.70039 | 3.32511 | 1.16816 | 5.89883 |
| 10 | 2.15585 | 2.55821 | 3.24697 | 3.94030 | 4.86518 | 6.73720 |
| 11 | 2.60321 | 3.15347 | 3.81575 | 4.57481 | 5.57779 | 7.58412 |
| 12 | 3.07382 | 3.57056 | 4.40379 | 5.22603 | 6.30380 | 8.43842 |
| 13 | 3.56503 | 4.10691 | 5.00874 | 5.89186 | 7.04150 | 9.29906 |
| 14 | 4.07468 | 4.66043 | 5.62872 | 6.57063 | 7.78953 | 10.1653 |
| 15 | 4.60094 | 5.22935 | 6.26214 | 7.26094 | 8.54675 | 11.0365 |
| 16 | 5.14224 | 5.81221 | 6.90766 | 7.96164 | 9.31223 | 11.9122 |
| 17 | 5.69724 | 6.40776 | 7.56418 | 8.67176 | 10.0852 | 12.7919 |
| 18 | 6.26481 | 7.01491 | 8.23075 | 9.39046 | 10.8649 | 13.6753 |
| 19 | 6.84398 | 7.63273 | 8.90655 | 10.1170 | 11.6509 | 14.5620 |
| 20 | 7.43386 | 8.26040 | 9.59083 | 10.8508 | 12.4426 | 15.4518 |
| 21 | 8.03366 | 8.89720 | 10.28293 | 11.5613 | 13.2396 | 16.3444 |
| 22 | 8.64272 | 9.54279 | 10.3923 | 12.3380 | 14.0415 | 17.2396 |
| 23 | 9.26042 | 10.19567 | 11.6885 | 13.0905 | 14.8479 | 18.1373 |
| 24 | 9.88623 | 10.8564 | 12.4011 | 13.8484 | 15.6587 | 19.0372 |
| 25 | 10.5197 | 11.5240 | 13.1197 | 14.6114 | 16.4734 | 19.9393 |
| 26 | 11.1603 | 12.1981 | 13.8439 | 15.3791 | 17.2919 | 20.8434 |
| 27 | 11.8076 | 12.8786 | 14.5733 | 16.1513 | 18.1138 | 21.7494 |
| 28 | 12.4613 | 13.5648 | 15.3079 | 16.9279 | 18.9392 | 22.6572 |
| 29 | 13.1211 | 14.2565 | 16.0471 | 17.7083 | 19.7677 | 23.5666 |
| 30 | 13.7867 | 14.9535 | 16.7908 | 18.4926 | 20.5992 | 24.4776 |
| 40 | 20.7065 | 22.1643 | 24.4331 | 26.5093 | 29.0505 | 33.6603 |
| 50 | 27.9907 | 29.7067 | 32.3574 | 34.7642 | 37.6886 | 42.9421 |
| 60 | 35.5346 | 37.4848 | 40.4817 | 43.1879 | 46.4589 | 52.2938 |
| 70 | 43.2752 | 45.4418 | 48.7576 | 51.7393 | 55.3290 | 61.6983 |
| 80 | 51.1720 | 53.5400 | 57.1532 | 60.3915 | 64.2778 | 71.1445 |
| 90 | 59.1963 | 61.7541 | 65.6466 | 69.1260 | 73.2912 | 80.6247 |
| 100 | 67.3276 | 70.0648 | 74.2219 | 77.9295 | 82.3581 | 90.1332 |

Table B.3 (continued)

| $\nu \backslash P$ | 0.500 | 0.250 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
|--------------------|----------|---------|---------|---------|---------|---------|---------|
| 1 | 0.454937 | 1.32330 | 2.70554 | 3.84146 | 5.02389 | 6.63490 | 7.87944 |
| 2 | 1.38629 | 2.77259 | 4.60517 | 5.99147 | 7.37776 | 9.21034 | 10.5966 |
| 3 | 2.36597 | 1.10835 | 6.25139 | 7.81473 | 9.34840 | 11.3449 | 12.8381 |
| 4 | 3.35670 | 5.38527 | 7.77944 | 9.48773 | 11.1433 | 13.2767 | 14.8602 |
| 5 | 4.35146 | 6.62568 | 9.23635 | 11.0705 | 12.8325 | 15.0863 | 16.7496 |
| 6 | 5.34812 | 7.84080 | 10.6446 | 12.5916 | 14.4494 | 16.8119 | 18.5476 |
| 7 | 6.34581 | 9.03715 | 12.0170 | 14.0671 | 16.0128 | 18.4753 | 20.2777 |
| 8 | 7.34412 | 10.2188 | 13.3616 | 15.5073 | 17.5346 | 20.0902 | 21.9550 |
| 9 | 8.34283 | 11.3887 | 14.6837 | 16.9190 | 19.0228 | 21.6660 | 23.5893 |
| 10 | 9.34182 | 12.5489 | 15.9871 | 18.3070 | 20.4831 | 23.2093 | 25.1882 |
| 11 | 10.3410 | 13.7007 | 17.2750 | 19.6751 | 21.9200 | 24.7250 | 26.7569 |
| 12 | 11.3403 | 14.8454 | 18.5494 | 21.0261 | 23.3367 | 26.2170 | 28.2995 |
| 13 | 12.3398 | 15.9839 | 19.8119 | 22.3621 | 24.7356 | 27.6883 | 29.8194 |
| 14 | 13.3393 | 17.1170 | 21.0642 | 23.6848 | 26.1190 | 29.1413 | 31.3193 |
| 15 | 14.3389 | 18.2451 | 22.3072 | 24.9958 | 27.4884 | 30.5779 | 32.8013 |
| 16 | 15.3385 | 19.3688 | 23.5418 | 26.2962 | 28.8454 | 31.9999 | 34.2672 |
| 17 | 16.3381 | 20.4887 | 24.4690 | 27.5871 | 30.1910 | 33.4087 | 35.7185 |
| 18 | 17.3379 | 21.6049 | 25.9894 | 28.8693 | 31.5264 | 34.8053 | 37.1564 |
| 19 | 18.3376 | 22.7178 | 27.2036 | 30.1435 | 32.8523 | 36.1908 | 38.5822 |
| 20 | 19.3374 | 23.8277 | 28.4120 | 31.4104 | 34.1696 | 37.5662 | 39.9968 |
| 21 | 20.3372 | 24.9348 | 29.6151 | 32.6705 | 35.4789 | 38.9321 | 41.4010 |
| 22 | 21.3370 | 26.0393 | 30.8133 | 33.9244 | 36.7807 | 40.2894 | 42.7956 |
| 23 | 22.3369 | 27.1413 | 32.0069 | 35.1725 | 38.0757 | 41.6384 | 44.1813 |
| 24 | 23.3367 | 28.2412 | 33.1963 | 36.4151 | 39.3641 | 42.9798 | 45.5585 |
| 25 | 24.3366 | 29.3389 | 34.3816 | 37.6525 | 40.6465 | 44.3141 | 46.9278 |
| 26 | 25.3364 | 30.4345 | 35.5631 | 38.8852 | 41.9232 | 45.6417 | 48.2899 |
| 27 | 26.3363 | 31.5284 | 36.7412 | 40.1133 | 43.1944 | 46.9630 | 49.6449 |
| 28 | 27.3363 | 32.6205 | 37.9159 | 41.3372 | 44.4607 | 48.2782 | 50.9933 |
| 29 | 28.3362 | 33.7109 | 39.0875 | 42.5569 | 45.7222 | 49.5879 | 52.3356 |
| 30 | 29.3360 | 34.7998 | 40.2560 | 43.7729 | 46.9792 | 50.8922 | 53.6720 |
| 40 | 39.3354 | 45.6160 | 51.8050 | 55.7585 | 59.3417 | 63.6907 | 66.7659 |
| 50 | 49.3349 | 56.3336 | 63.1671 | 67.5048 | 71.4202 | 76.1539 | 79.4900 |
| 60 | 59.3347 | 66.9814 | 74.3970 | 79.0819 | 83.2976 | 88.3794 | 91.9517 |
| 70 | 69.3344 | 77.5766 | 85.5271 | 90.5312 | 95.0231 | 100.425 | 104.215 |
| 80 | 79.3343 | 88.1303 | 96.5782 | 101.879 | 106.629 | 112.329 | 116.321 |
| 90 | 89.3342 | 98.6499 | 107.565 | 113.145 | 118.136 | 124.116 | 128.299 |
| 100 | 99.3341 | 109.141 | 118.498 | 124.342 | 129.561 | 135.807 | 140.169 |

(Source: Reprinted, by permission of the Biometrika Trustees from C. M. Thompson, "Tables of Percentage Points of the χ^2 Distribution," *Biometrika* 32 (1941): 188–189.)

Table B.4 *F*-Distribution: 5% Points

The first column lists the number of denominator degrees of freedom (v_2). The headings of the other columns list the numerator degrees of freedom (v_1). The table entry is the value of c for which $P(F_{v_1, v_2} \geq c) = .05$.

| $v_2 \backslash v_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1 | 161.45 | 199.50 | 215.71 | 224.58 | 230.16 | 233.99 | 236.77 | 238.88 | 240.54 |
| 2 | 18.513 | 19.000 | 19.164 | 19.247 | 19.296 | 19.330 | 19.353 | 19.371 | 19.385 |
| 3 | 10.128 | 9.5521 | 9.2766 | 9.1172 | 9.0135 | 8.9406 | 8.8868 | 8.8452 | 8.8123 |
| 4 | 7.7086 | 6.9443 | 6.5914 | 6.3883 | 6.3560 | 6.1631 | 6.0942 | 6.0410 | 5.9988 |
| 5 | 6.6079 | 5.7861 | 5.4095 | 5.1922 | 5.0503 | 4.9503 | 4.8759 | 4.8183 | 4.7725 |
| 6 | 5.9874 | 5.1433 | 4.7571 | 4.5337 | 4.3874 | 4.2839 | 4.2066 | 4.1468 | 4.0990 |
| 7 | 5.5914 | 4.7374 | 4.3468 | 4.1203 | 3.9715 | 3.8660 | 3.7870 | 3.7257 | 3.6767 |
| 8 | 5.3177 | 4.4590 | 4.0662 | 3.8378 | 3.6875 | 3.5806 | 3.5005 | 3.4381 | 3.3881 |
| 9 | 5.1174 | 4.2565 | 3.8626 | 3.6331 | 3.4817 | 3.3738 | 3.2927 | 3.2296 | 3.1789 |
| 10 | 4.9646 | 4.1028 | 3.7083 | 3.4780 | 3.3258 | 3.2172 | 3.1355 | 3.0717 | 3.0204 |
| 11 | 4.8443 | 3.9823 | 3.5874 | 3.3567 | 3.2039 | 3.0946 | 3.0123 | 2.9480 | 2.8962 |
| 12 | 4.7472 | 3.8856 | 3.4903 | 3.2592 | 3.1059 | 2.9961 | 3.9134 | 2.8486 | 2.7964 |
| 13 | 4.6672 | 3.8056 | 3.4105 | 3.1791 | 3.0254 | 2.9153 | 2.8321 | 2.7669 | 2.7144 |
| 14 | 4.6001 | 3.7389 | 3.3439 | 3.1122 | 2.9582 | 2.8477 | 2.7642 | 2.6987 | 2.6458 |
| 15 | 4.5431 | 3.6823 | 3.2874 | 3.0556 | 2.9013 | 2.7905 | 2.7066 | 2.6408 | 2.5876 |
| 16 | 4.4940 | 3.6337 | 3.2389 | 3.0069 | 2.8524 | 2.7413 | 2.6572 | 2.5911 | 2.5377 |
| 17 | 4.4513 | 3.5915 | 3.1968 | 2.9647 | 2.8100 | 2.6987 | 2.6143 | 2.5480 | 2.4943 |
| 18 | 4.4139 | 3.5546 | 3.1599 | 2.9277 | 2.7729 | 2.6613 | 2.5767 | 2.5102 | 2.4563 |
| 19 | 4.3808 | 3.5219 | 3.1274 | 2.8951 | 2.7401 | 2.6283 | 2.5435 | 2.4768 | 2.4227 |
| 20 | 4.3513 | 3.4928 | 3.0984 | 2.8661 | 2.7109 | 2.5990 | 2.5140 | 2.4471 | 2.3928 |
| 21 | 4.3248 | 3.4668 | 3.0725 | 2.8401 | 2.6848 | 2.5727 | 2.4876 | 2.4205 | 2.3661 |
| 22 | 4.3009 | 3.4434 | 3.0491 | 2.8167 | 2.6613 | 2.5491 | 2.4638 | 2.3965 | 2.3419 |
| 23 | 4.2793 | 3.4221 | 3.0280 | 2.7955 | 2.6400 | 2.5277 | 2.4422 | 2.3748 | 2.3201 |
| 24 | 4.2597 | 3.4028 | 3.0088 | 3.7763 | 2.6207 | 2.5082 | 2.4226 | 2.3551 | 2.3002 |
| 25 | 4.2417 | 3.3852 | 2.9912 | 2.7587 | 2.6030 | 2.4904 | 2.4047 | 2.3371 | 2.2821 |
| 26 | 4.2252 | 3.3690 | 2.9751 | 2.7426 | 2.5868 | 2.4741 | 2.3883 | 2.3205 | 2.2655 |
| 27 | 4.2100 | 3.3541 | 2.9604 | 2.7278 | 2.5719 | 2.4591 | 2.3732 | 2.3053 | 2.2501 |
| 28 | 4.1960 | 3.3404 | 2.9467 | 2.7141 | 2.5581 | 2.4453 | 2.3593 | 2.2913 | 2.2360 |
| 29 | 4.1830 | 3.3277 | 2.9340 | 2.7014 | 2.5454 | 2.4324 | 2.3463 | 2.2782 | 2.2229 |
| 30 | 4.1709 | 3.3158 | 2.9223 | 2.6896 | 2.5336 | 2.4205 | 2.3343 | 2.2662 | 2.2107 |
| 40 | 4.0848 | 3.2317 | 2.8387 | 2.6060 | 2.4495 | 2.3359 | 2.2490 | 2.1802 | 2.1240 |
| 60 | 4.0012 | 3.1504 | 2.7581 | 2.5252 | 2.3683 | 2.2540 | 2.1665 | 2.0970 | 2.0401 |
| 120 | 3.9201 | 3.0718 | 2.6802 | 2.4472 | 2.2900 | 2.1750 | 2.0867 | 2.0164 | 1.9588 |
| ∞ | 3.8415 | 2.9957 | 2.6049 | 2.3719 | 2.2141 | 2.0986 | 2.0096 | 1.9354 | 1.8799 |

Table B.4 F-Distribution: 5% Points (*continued*)

| $\frac{V_1}{V_2}$ | 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ |
|-------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|----------|
| 1 | 241.88 | 243.91 | 245.95 | 248.01 | 249.05 | 250.09 | 251.14 | 252.20 | 253.25 | 254.32 |
| 2 | 19.396 | 19.413 | 19.429 | 19.446 | 19.454 | 19.462 | 19.471 | 19.479 | 19.487 | 19.496 |
| 3 | 8.7855 | 8.7446 | 8.7029 | 8.6602 | 8.6385 | 8.6166 | 8.5944 | 8.5720 | 8.5494 | 8.5265 |
| 4 | 5.9644 | 5.9117 | 5.8578 | 5.8025 | 5.7744 | 5.7459 | 5.7170 | 5.6878 | 5.6581 | 5.6281 |
| 5 | 4.7351 | 4.6777 | 4.6188 | 4.5581 | 4.5272 | 4.4957 | 4.4638 | 4.4314 | 4.3984 | 4.3650 |
| 6 | 4.0600 | 3.9999 | 3.9381 | 3.8742 | 3.8415 | 3.8082 | 3.7743 | 3.7398 | 3.7047 | 3.6688 |
| 7 | 3.6365 | 3.5747 | 3.5108 | 3.4445 | 3.4105 | 3.3758 | 3.3404 | 3.3043 | 3.2674 | 3.2298 |
| 8 | 3.3472 | 3.2840 | 3.2184 | 3.1503 | 3.1152 | 3.0794 | 3.0428 | 3.0053 | 2.9669 | 2.9276 |
| 9 | 3.1373 | 3.0729 | 3.0061 | 2.9365 | 2.9005 | 2.8637 | 2.8259 | 2.7872 | 2.7475 | 2.7067 |
| 10 | 2.9782 | 2.9130 | 2.8450 | 2.7740 | 2.7372 | 2.6996 | 2.6609 | 2.6211 | 2.5801 | 2.5379 |
| 11 | 2.8536 | 2.7876 | 2.7186 | 2.6464 | 2.6090 | 2.5705 | 2.5309 | 2.4901 | 2.4480 | 2.4045 |
| 12 | 2.7534 | 2.6866 | 2.6169 | 2.5436 | 2.5055 | 2.4663 | 2.4259 | 2.3842 | 2.3410 | 2.2962 |
| 13 | 2.6710 | 2.6037 | 2.5331 | 2.4589 | 2.4202 | 2.3803 | 2.3392 | 2.2966 | 2.2524 | 2.2064 |
| 14 | 2.6021 | 2.5342 | 2.4630 | 2.3879 | 2.3487 | 2.3082 | 2.2664 | 2.2230 | 2.1778 | 2.1307 |
| 15 | 2.5437 | 2.4753 | 2.4035 | 2.3275 | 2.2878 | 2.2468 | 2.2043 | 2.1601 | 2.1141 | 2.0658 |
| 16 | 2.4935 | 2.4247 | 2.3522 | 2.2756 | 2.2354 | 2.1938 | 2.1507 | 2.1058 | 2.0589 | 2.0096 |
| 17 | 2.4499 | 2.3807 | 2.3077 | 2.2304 | 2.1898 | 2.1477 | 2.1040 | 2.0584 | 2.0107 | 1.9604 |
| 18 | 2.4117 | 2.3421 | 2.2686 | 2.1906 | 2.1497 | 2.1071 | 2.0629 | 2.0166 | 1.9681 | 1.9168 |
| 19 | 2.3779 | 2.3080 | 2.2341 | 2.1555 | 2.1141 | 2.0712 | 2.0264 | 1.9796 | 1.9302 | 1.8780 |
| 20 | 2.3479 | 2.2776 | 2.2033 | 2.1242 | 2.0825 | 2.0391 | 1.9938 | 1.9464 | 1.8963 | 1.8432 |
| 21 | 2.3210 | 2.2504 | 2.1757 | 2.0960 | 2.0540 | 2.0102 | 1.9645 | 1.9165 | 1.8657 | 1.8117 |
| 22 | 2.2967 | 2.2258 | 2.1508 | 2.0707 | 2.0283 | 1.9842 | 1.9380 | 1.8895 | 1.8380 | 1.7831 |
| 23 | 2.2747 | 2.2036 | 2.1282 | 1.0476 | 2.0050 | 1.9605 | 1.9139 | 1.8649 | 1.8128 | 1.7570 |
| 24 | 2.2547 | 2.1834 | 2.1077 | 2.0267 | 1.9838 | 1.9390 | 1.8920 | 1.8424 | 1.7897 | 1.7331 |
| 25 | 2.2365 | 2.1649 | 2.0889 | 2.0075 | 1.9643 | 1.9192 | 1.8718 | 1.8217 | 1.7684 | 1.7110 |
| 26 | 2.2197 | 2.1479 | 2.0716 | 1.9898 | 1.9464 | 1.9010 | 1.8533 | 1.8027 | 1.7488 | 1.6906 |
| 27 | 2.2043 | 2.1323 | 2.0558 | 1.9736 | 1.9299 | 1.8842 | 1.8361 | 1.7851 | 1.7307 | 1.6717 |
| 28 | 2.1900 | 2.1179 | 2.0411 | 1.9586 | 1.9147 | 1.8687 | 1.8203 | 1.7689 | 1.7138 | 1.6541 |
| 29 | 2.1768 | 2.1045 | 2.0245 | 1.9446 | 1.9005 | 1.8543 | 1.8055 | 1.7537 | 1.6981 | 1.6377 |
| 30 | 2.1646 | 2.0921 | 2.0148 | 1.9317 | 1.8874 | 1.8409 | 1.7918 | 1.7396 | 1.6835 | 1.6223 |
| 40 | 2.0772 | 2.0035 | 1.9245 | 1.8389 | 1.7929 | 1.7444 | 1.6928 | 1.6373 | 1.5766 | 1.5089 |
| 60 | 1.9926 | 1.9174 | 1.8364 | 1.7480 | 1.7001 | 1.6491 | 1.5943 | 1.5343 | 1.4673 | 1.3893 |
| 120 | 1.9105 | 1.8337 | 1.7505 | 1.6587 | 1.6084 | 1.5543 | 1.4952 | 1.4290 | 1.3519 | 1.2539 |
| ∞ | 1.8307 | 1.7522 | 1.6664 | 1.5705 | 1.5173 | 1.4591 | 1.3940 | 1.3180 | 1.2214 | 1.0000 |

(Source: Reprinted, by permission of the Biometrika Trustees from M. Merrington and C. M. Thompson, "Tables of Percentage Points of the Inverted Beta(F) Distribution," *Biometrika* 33 (1943): 80-81.)

Table B.5 *F*-Distribution: 1% Points

The first column lists the number of denominator degrees of freedom (v_2). The headings of the other columns list the numerator degrees of freedom (v_1). The table entry is the value of c for which $P(F_{v_1, v_2} \geq c) = .01$.

| $v_1 \backslash v_2$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1 | 4052.2 | 4999.5 | 5403.3 | 5624.6 | 5763.7 | 5859.0 | 5928.3 | 5981.6 | 6022.5 |
| 2 | 98.503 | 99.000 | 99.166 | 99.249 | 99.299 | 99.332 | 99.356 | 99.374 | 99.388 |
| 3 | 34.116 | 30.817 | 29.457 | 28.710 | 28.237 | 27.911 | 27.672 | 27.489 | 27.345 |
| 4 | 21.198 | 18.000 | 16.694 | 15.977 | 15.522 | 15.207 | 14.976 | 14.799 | 14.659 |
| 5 | 16.258 | 13.274 | 12.060 | 11.392 | 10.967 | 10.672 | 10.456 | 10.289 | 10.158 |
| 6 | 13.745 | 10.925 | 9.7795 | 9.1483 | 8.7459 | 8.4661 | 8.2600 | 8.1016 | 7.9761 |
| 7 | 12.246 | 9.5466 | 8.4513 | 7.8467 | 7.4604 | 7.1914 | 6.9928 | 6.8401 | 6.7188 |
| 8 | 11.259 | 8.6491 | 7.5910 | 7.0060 | 6.6318 | 6.3707 | 6.1776 | 6.0289 | 5.9106 |
| 9 | 10.561 | 8.0215 | 6.9919 | 6.4221 | 6.0569 | 5.8018 | 5.6129 | 5.4671 | 5.3511 |
| 10 | 10.044 | 7.5594 | 6.5523 | 5.9943 | 5.6363 | 5.3858 | 5.2001 | 5.0567 | 4.9424 |
| 11 | 9.6460 | 7.2057 | 6.2167 | 5.6683 | 5.3160 | 5.0692 | 4.8861 | 4.7445 | 4.6315 |
| 12 | 9.3302 | 6.9266 | 5.9526 | 5.4119 | 5.0643 | 4.8206 | 4.6395 | 4.4994 | 4.3875 |
| 13 | 9.0738 | 6.7010 | 5.7394 | 5.2053 | 4.8616 | 4.6204 | 4.4410 | 4.3021 | 4.1911 |
| 14 | 8.8616 | 6.5149 | 5.5639 | 5.0354 | 4.6950 | 4.4558 | 4.2779 | 4.1399 | 4.0297 |
| 15 | 8.6831 | 6.3589 | 5.4170 | 4.8932 | 4.5556 | 4.3183 | 4.1415 | 4.0045 | 3.8948 |
| 16 | 8.5310 | 6.2262 | 5.2922 | 4.7726 | 4.4374 | 4.2016 | 4.0259 | 3.8896 | 3.7804 |
| 17 | 8.3997 | 6.1121 | 5.1850 | 4.6690 | 4.3359 | 4.1015 | 3.9267 | 3.7910 | 3.6822 |
| 18 | 8.2854 | 6.0129 | 5.0919 | 4.5790 | 4.2479 | 4.0146 | 3.8406 | 3.7054 | 3.5971 |
| 19 | 8.1850 | 5.9259 | 5.0103 | 4.5003 | 4.1708 | 3.9386 | 3.7653 | 3.6305 | 3.5225 |
| 20 | 8.0906 | 5.8489 | 4.9382 | 4.4307 | 4.1027 | 3.8714 | 3.6987 | 3.5644 | 3.4567 |
| 21 | 8.0166 | 5.7804 | 4.8740 | 4.3688 | 4.0421 | 3.8117 | 3.6396 | 3.5056 | 3.3981 |
| 22 | 7.9454 | 5.7190 | 4.8166 | 4.3134 | 3.9880 | 3.7583 | 3.5867 | 3.4530 | 3.3458 |
| 23 | 7.8811 | 5.6637 | 4.7649 | 4.2635 | 3.9392 | 3.7102 | 3.5390 | 3.4057 | 3.2986 |
| 24 | 7.8229 | 5.6131 | 4.7181 | 4.2184 | 3.8951 | 3.6667 | 3.4959 | 3.3629 | 3.2560 |
| 25 | 7.7689 | 5.5680 | 4.6755 | 4.1774 | 3.8550 | 3.6272 | 3.4568 | 3.3239 | 3.2172 |
| 26 | 7.7213 | 5.5263 | 4.6366 | 4.1400 | 3.8183 | 3.5911 | 3.4210 | 3.2884 | 3.1818 |
| 27 | 7.6767 | 5.4881 | 4.6009 | 4.1056 | 3.7848 | 3.5580 | 3.3882 | 3.2558 | 3.1494 |
| 28 | 7.6356 | 5.4529 | 4.5681 | 4.0740 | 3.7539 | 3.5276 | 3.3581 | 3.2259 | 3.1195 |
| 29 | 7.5976 | 5.4205 | 4.5378 | 4.0449 | 3.7254 | 3.4995 | 3.3302 | 3.1982 | 3.0920 |
| 30 | 7.5625 | 5.3904 | 4.5097 | 4.0179 | 3.6990 | 3.4735 | 3.3045 | 3.1726 | 3.0665 |
| 40 | 7.3141 | 5.1785 | 4.3126 | 3.8283 | 3.5138 | 3.2910 | 3.1238 | 2.9930 | 2.8876 |
| 60 | 7.0771 | 4.9774 | 4.1259 | 3.6491 | 3.3389 | 3.1187 | 2.9530 | 2.8233 | 2.7185 |
| 120 | 6.8510 | 4.7865 | 3.9493 | 3.4796 | 3.1735 | 2.9559 | 2.7918 | 2.6629 | 2.5586 |
| ∞ | 6.6349 | 4.6052 | 3.7816 | 3.3192 | 3.0173 | 2.8020 | 2.6393 | 2.5113 | 2.4073 |

Table B.5 *F*-Distribution: 1% Points (*continued*)

| $v_1 \backslash v_2$ | 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ |
|----------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|----------|
| 1 | 6055.8 | 6106.3 | 6157.3 | 6208.7 | 6234.6 | 6260.7 | 6286.8 | 6313.0 | 6339.4 | 6366.0 |
| 2 | 99.399 | 99.416 | 99.449 | 99.458 | 99.458 | 99.466 | 99.474 | 99.483 | 99.491 | 99.501 |
| 3 | 27.229 | 27.052 | 26.872 | 26.690 | 26.598 | 26.505 | 26.411 | 26.316 | 26.221 | 26.125 |
| 4 | 14.546 | 14.374 | 14.198 | 14.020 | 13.929 | 13.838 | 13.745 | 13.652 | 13.558 | 13.463 |
| 5 | 10.051 | 9.8883 | 9.7222 | 9.5527 | 9.4665 | 9.3793 | 9.2912 | 9.2020 | 9.1118 | 9.0204 |
| 6 | 7.8741 | 7.7183 | 7.5590 | 7.3958 | 7.3127 | 7.2285 | 7.1432 | 7.0568 | 6.9690 | 6.8801 |
| 7 | 6.6201 | 6.4691 | 6.3143 | 6.1554 | 6.0743 | 5.9921 | 5.9084 | 5.8236 | 5.7372 | 5.6495 |
| 8 | 5.8143 | 5.6668 | 5.5151 | 5.3591 | 5.2793 | 5.1981 | 5.1156 | 5.0316 | 4.9460 | 4.8588 |
| 9 | 5.2565 | 5.1114 | 4.9621 | 4.8080 | 4.7290 | 4.6486 | 4.5667 | 4.4831 | 4.3978 | 4.3105 |
| 10 | 4.8492 | 4.7059 | 4.5582 | 4.4054 | 4.3269 | 4.2469 | 4.1653 | 4.0819 | 3.9965 | 3.9090 |
| 11 | 4.5393 | 4.3974 | 4.2509 | 4.0990 | 4.0209 | 3.9411 | 3.8596 | 3.7761 | 3.6904 | 3.6025 |
| 12 | 4.2961 | 4.1553 | 4.0096 | 3.8584 | 3.7805 | 3.7008 | 3.6192 | 3.5355 | 3.4494 | 3.3608 |
| 13 | 4.1003 | 3.9603 | 3.8154 | 3.6646 | 3.5868 | 3.5070 | 3.4253 | 3.3413 | 3.2548 | 3.1654 |
| 14 | 3.9394 | 3.8001 | 3.6557 | 3.5052 | 3.4274 | 3.3476 | 3.2656 | 3.1813 | 3.0942 | 3.0040 |
| 15 | 3.8049 | 3.6662 | 3.5255 | 3.3719 | 3.2940 | 3.2141 | 3.1319 | 3.0471 | 2.9595 | 2.8684 |
| 16 | 3.6909 | 3.5527 | 3.4089 | 3.2588 | 3.1808 | 3.1007 | 3.0182 | 2.9330 | 2.8447 | 2.7528 |
| 17 | 3.5931 | 3.4552 | 3.3117 | 3.1615 | 3.0835 | 3.0032 | 2.9205 | 2.8348 | 2.7459 | 2.6530 |
| 18 | 3.5082 | 3.3706 | 3.2273 | 3.0771 | 2.9990 | 2.9185 | 2.8354 | 2.7493 | 2.6597 | 2.5660 |
| 19 | 3.4338 | 3.2965 | 3.1533 | 3.0031 | 2.9249 | 2.8442 | 2.7608 | 2.6742 | 2.5839 | 2.4893 |
| 20 | 3.3682 | 3.2311 | 3.0880 | 2.9377 | 2.8594 | 2.7785 | 2.6947 | 2.6077 | 2.5168 | 2.4212 |
| 21 | 3.3098 | 3.1729 | 3.0299 | 2.8796 | 2.8011 | 2.7200 | 2.6359 | 2.5484 | 2.4568 | 2.3603 |
| 22 | 3.2576 | 3.1209 | 2.9780 | 2.8274 | 2.7488 | 2.6675 | 2.5831 | 2.4951 | 2.4029 | 2.3055 |
| 23 | 3.2106 | 3.0740 | 2.9311 | 2.7805 | 2.7017 | 2.6202 | 2.5355 | 2.4471 | 2.3542 | 2.2559 |
| 24 | 3.1681 | 3.0316 | 2.8887 | 2.7380 | 2.6591 | 2.5773 | 2.4923 | 2.4035 | 2.3099 | 2.2107 |
| 25 | 3.1294 | 2.9331 | 2.8502 | 2.6993 | 2.6203 | 2.5383 | 2.4530 | 2.3667 | 2.2695 | 2.1694 |
| 26 | 3.0941 | 2.9576 | 2.8150 | 2.6640 | 2.5848 | 2.5026 | 2.4170 | 2.3273 | 2.2325 | 2.1315 |
| 27 | 3.0618 | 2.2956 | 2.7827 | 2.6316 | 2.5522 | 2.4699 | 2.3840 | 2.2938 | 2.1984 | 2.0965 |
| 28 | 3.0320 | 2.8959 | 2.7530 | 2.6017 | 2.5223 | 2.4397 | 2.3535 | 2.2629 | 2.1670 | 2.0642 |
| 29 | 3.0045 | 2.8685 | 2.7256 | 2.5742 | 2.4946 | 2.4118 | 2.3253 | 2.2344 | 2.1378 | 2.0342 |
| 30 | 2.9791 | 2.8431 | 2.7002 | 2.5487 | 2.4689 | 2.3680 | 2.2992 | 2.2079 | 2.1107 | 2.0062 |
| 40 | 2.8005 | 2.6648 | 2.5216 | 2.3689 | 2.2880 | 2.2034 | 2.1142 | 2.0194 | 1.9172 | 1.8047 |
| 60 | 2.6318 | 2.4961 | 2.3523 | 2.1978 | 2.1154 | 2.0285 | 1.9360 | 1.8363 | 1.7263 | 1.6006 |
| 120 | 2.4721 | 2.3363 | 2.1915 | 2.0346 | 1.9500 | 1.8600 | 1.7628 | 1.6557 | 1.5330 | 1.3805 |
| ∞ | 2.3209 | 2.1848 | 2.0385 | 1.8783 | 1.7908 | 1.6964 | 1.5923 | 1.4730 | 1.3246 | 1.0000 |

(Source: Reprinted, by permission of the Biometrika Trustees from M. Merrington and C. M. Thompson, "Tables of Percentage Points of the Inverted Beta(*F*) Distribution," *Biometrika* 33 (1943): 84–85.)

Index

A

abbreviated set definition for events, 60
absolute convergence, 111, 113
acceptance region, 515
admissible estimator, 375
with respect to SMSE, 377
admissibility of statistical test, 535
almost-sure convergence (or convergence with probability 1), 253
implies convergence in probability, 255
of continuous functions of X_n , 256
of relative frequency, 266
almost-sure limit, 253
alternative hypothesis, 527
approximating χ^2 probabilities via the normal distribution, 273

approximating binomial probabilities via the normal distribution, 272
via the Poisson distribution, 178
approximation to critical values of the χ^2 -distribution, 674
 \arg , 466
 argmax , 465
 argmin , 433
asymptotic admissibility, 386
asymptotic confidence interval, 646
asymptotic distribution, 239, 275
of differentiable functions of asymptotically normally distributed random variables, 286–287
of $g(X_n, \Theta_n)$, 239
relationship with limiting distribution, 239, 272
asymptotic efficiency or best asymptotically normal (BAN), 388, 419

asymptotic estimator properties, nonuniqueness, 383
asymptotic mean square error, 383
asymptotic normality, of GMM Estimator, 500
of MOM Estimator of Θ , 493
of the least-squares estimator (scalar β case), 275, 280
of sample moments about the origin, 318
of $\hat{\beta}$ in the GLM, 443–447
(in the GLM, 447–448)
asymptotic nonpositively correlated sequence, 266
asymptotic pivotal quantity, 650
asymptotic relative efficiency (ARE), 384, 386
asymptotic tests, 526
asymptotically equivalent tests, 614
asymptotically relatively more efficient estimator, 384

- asymptotically unbiased, estimator, 386
test, 609
- asymptotically uniformly most powerful, 600
- at most of order n^k , 231
- at most of order n^k in probability, 248
- autocorrelation, 432, 453
- autocorrelation process (first-order), 267
- autocovariance, 432
- auxiliary random variable in change of variables approach, 340, 358
- axiomatic, 8
- axiomatic foundation for probability theory, 13
- B**
- Bayesian approach to inference, 369
- Bayes rule, 34
- Behrens-Fisher problem, 669
- Bernoulli density, 172
- Bernoulli trials, 173, 176
- best linear prediction of Y outcome, 155
- best linear unbiased estimator (BLUE) or minimum variance linear unbiased estimator (MVBLUE), 382
- BLUE or MVBLUE of population mean, 382
- beta density, 195, 196
- beta function, 195
- bias, 374
- bias-adjusted maximum likelihood estimator, 471
- bias matrix, 375
- bias vector, 374
- biased, 378
- binary relation, 687
- binomial density, 173, 196
- binomial theorem, 137, 696
- Bonferroni joint tests of $\mathbf{R}_i\boldsymbol{\beta} = \mathbf{r}_i$,
 $\mathbf{R}_i\boldsymbol{\beta} \leq \mathbf{r}_i$, or $\mathbf{R}_i\boldsymbol{\beta} \geq \mathbf{r}_i$,
 $i = 1, \dots, m$, 631
- Bonferroni's inequality, 17
- Boole's inequality, 16
- Borel sets, 21
- bounded in probability, 249
- bounded sequence of real numbers, 225
- C**
- Cartesian product, 688
- Cauchy's criterion for almost-sure convergence, 257
- Cauchy-Schwarz inequality, 150
- central limit theorem (CLT), 268
- for bounded m -dependent sequences, 281
- for bounded random variables, 275
- for independent bounded random vectors, 285
- for triangular arrays, 279
- Liapounov, 276, 279
- Lindberg, 274
- Lindberg-Levy, 270, 283
- Van Beeck's bound on approximation error, 273–274
- Zolotarev's bound on approximation error, 278
- certain event, 9
- change of variables technique, 335
- when $g(X)$ is multivariate and invertible, 340
- when $g(X)$ is univariate and invertible, 335
- when $g(X)$ is univariate and piecewise invertible, 338
- characteristic function, 145
- Chebyshev's Inequality, 133–134
- chi-square density, 192
- additivity property, 193
- from sums of squares of independent standard normal random variables, 200
- Cholesky decomposition, 207
- class of continuous density functions, 55, 69
- of discrete density functions, 55, 69
- classical assumptions of the general linear model, 430
- classical probability, 5, 18
- closed interval, 687
- rectangle, 21
- coefficient of determination, 434
- complement, 682
- complete sufficient statistics, 400
- in the exponential class of densities, 402
- component experiments, 305
- composite experiment, 305
- composite hypothesis, 510
- composite random variable, 48, 77
- computer generation of random-variable outcomes, 186
- conditional cumulative distribution function, 89
- conditional densities when $X \sim N(\mu, \Sigma)$, 209
- conditional expectation, 126, 131
- conditional probability, 23
- conditional probability density functions, 83, 88
- conditional sample space, 22
- confidence coefficient, 641–642
- confidence interval, 356
- asymptotic, for p of Bernoulli distribution, 646
- for Θ in an exponential distribution, 645
- for μ in a Normal Distribution, 643
- with confidence level γ , 641
- confidence limits, 641
- confidence region, 642
- duality with hypothesis tests, 654
- consistent asymptotically normal (CAN) estimators, 385
- consistent confidence region sequence, 649
- consistent estimator, 384
- MOM Estimator of Θ , 493
- MOM Estimator of $q(\Theta)$, 495
- MLE for Exponential Distribution, 476
- MLE for Gamma Distribution, 478
- MLE in GLM with $\mathbf{Y} \sim N(\mathbf{x}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$, 479
- \hat{S}^2 in the GLM, 441–442
- GMM Estimator, 498
- $\hat{\boldsymbol{\beta}}$, in the GLM, 439
- consistent sequence of tests, 535
- GLR test, 608
- LM test, 617

Wald test, 623
 consistent uniformly asymptotically normal (CUAN) class, 389
 continuity correction, 272
 continuous function, 229
 continuous probability density function, 52
 continuous random variable, 52
 continuous sample space, 3
 continuous uniform density, 186
 control charting, 592
 converge to y , 224
 convergence in distribution (CDFs), 234
 in terms of MGFs, 238
 in terms of PDFs, 235
 to a constant, 235
 convergence in mean square (or convergence in quadratic mean), 249
 implies convergence in probability, 252
 necessary and sufficient conditions for, 250
 convergence in probability, 241, 246
 for continuous functions, 244
 implies convergence in distribution, 246
 of relative frequency, 260
 properties, 245
 convergence mode relationships, 258
 convergence of a function sequence, 230
 convergence of a sequence of PDFs, 235
 convergent, 230
 converges almost surely, 253
 coordinate function, 67
 correlation, 151
 correlation bounds, 150
 and linearity between random variables, 152
 correlation matrix, 160–161
 countable additivity, 12
 countable set, 681
 countably infinite set, 681
 covariance, 149
 covariance bound, 150

covariance matrix, 157
 Cramér-Rao Lower Bound (CRLB), 412
 alternative form, 413
 attainment of the CRLB, 417
 for scalar function case, 415
 regularity conditions, 409
 Cramér-Wold device, 282
 for normal limiting distributions, 283
 critical (or rejection) region of statistical test, 515
 duality with confidence region, 643
 of test statistic, 526
 cumulant-generating function, 146
 univariate, 146
 multivariate, 147
 cumulants, 146
 cumulative distribution function (CDF), univariate, 60–61
 duality between CDFs and PDFs, 65, 73
 CDF approach for deriving PDF of $g(X)$, 333
 CDF properties, 63
 multivariate or joint CDF, 73

D

degenerate density function, 153
 random variable, 153
 degrees of freedom, 192, 200
 dependent events, 29
 dependent variable, 429
 design matrix, 431
 discrete probability density function, 50
 discrete random variable, 50
 discrete sample space, 3
 discrete uniform density, 170
 disjoint events, 4
 distinct PDFs, 368
 disturbance vector in the GLM, 429
 divergent sequence, 224
 domain and range of a function, 691
 double array of random variables, 279
 double expectation theorem, 127–128, 131

double exponential family of PDFs, 216
 duality between confidence and critical regions, 643

E

econometrics, 460
 efficient estimator, 380, 382
 element (or member of the sample space), 3
 elementary event, 4
 elements of a set, 679
 ellipsoid confidence region for β in the GLM, 647
 empirical distribution function (EDF), 308–309, 313–314
 empirical substitution principle, 314
 empty or null set, 682
 equivalent events, 46
 approach for finding density of $g(X)$, 334
 error vector in the GLM, 429
 errors in variables problem in the GLM, 459
 estimand, 370
 estimate of Θ or $q(\Theta)$, 370
 estimating Θ or $q(\Theta)$, 370
 estimator for Θ or $q(\Theta)$, 370
 estimator properties, 370
 admissibility, 375
 asymptotic admissibility, 386
 asymptotically efficient, 388, 419
 asymptotically relatively more efficient, 386
 asymptotically unbiased, 386
 best asymptotically normal (BAN), 419
 best linear unbiased estimator (BLUE), 382
 bias, 374
 consistent, 384
 consistent asymptotically normal (CAN), 385
 minimum variance linear unbiased estimator (MVLUE), 382
 minimum variance unbiased estimator (MVUE), 380

estimator properties (*cont.*)
 relatively more efficient, 377
 strong mean square error (SMSE)
 superior, 376
 unbiased, 378
 event, 3
 elementary, 4
 mutually exclusive or disjoint events, 4
 occurrence of, 4
 occurs with probability one, 54
 occurs with probability zero, 54
 event space, Υ , 9
 $\exp(n)$, 143
 expected value, 111, 113
 conditional, 126, 131
 existence of $E(X)$ does not imply that $Eg(X)$ exists, 120
 existence of $E(X)$ for bounded $R(X)$, 115
 of function of X , $Y = g(X)$, 117
 of a function of a multivariate random variable, 122
 of a matrix of random variables, 123
 of a product, 124, 149
 of a sum, 124
 properties, 121–124
 experiment, 2
 exponential density, 190
 memoryless property, 191
 exponential class of densities, 214
 and complete sufficient statistics, 402
 and sufficient statistics, 398
 differentiability of power function 558, 573
 natural parameterization, 573
 extreme values, 355

F

$f(x)$, 691
 F -density, 344
 fair or equitable game of chance, 165
 family of densities, 358
 finite additivity, 12
 finite set, 681
 for every or for all, \forall , 691

function f from A to B ($f : A \rightarrow B$), 691
 functions of random variables,
 deriving probability densities, 331–332
 change of variables approach, 335–341
 CDF approach, 332–333
 equivalent events approach (discrete case), 334
 MGF approach, 332

G

gamma density, 187
 additivity, 189
 effect of scaling, 189
 inverse additivity, 190
 wear-out effect, 191
 work-hardening effect, 191
 gamma function, 187
 general linear model (GLM), 428
 classical assumptions of, 430
 violations of classic GLM assumptions, 452
 generalized likelihood ratio (GLR), 601
 generalized likelihood ratio (GLR) test, 601
 asymptotic distribution when $H_0 : R(\Theta) = r$ is true, 610
 asymptotic distribution when $H_a : R(\Theta) \neq r$ is true for local alternatives, 613
 asymptotic power of GLR Test of $H_0 : \Theta = \Theta_0$ in Exponential Population, 615
 equivalence with Neyman-Pearson most powerful test for simple hypotheses, 602
 generalized least-squares estimator of β in GLM, 456, 508
 generalized method of moments (GMM) estimator, 497–498
 geometric density, 176
 memoryless property, 176
 geometric mean, 591, 620
 Glivenko-Cantelli theorem, 313
 goodness-of-fit test (χ^2), 655

H

heteroskedasticity, 453
 homoskedasticity, 432
 hypergeometric density, 183, 185
 hypothesis, 509
 abbreviated representation of, 513
 one-sided alternative hypothesis, 549
 rejected at the α -level of significance, 533
 rejected using a size- α test, 533
 statistical, 510
 two-sided alternative hypothesis, 550
 hypothesis testing (see testing)

I

ideal statistical test, 517, 531
 iff (if and only if), 680
 iid (independent and identically distributed), 223
 Wald-Wolfowitz (WW) runs test of, 663
 image of x under f , 689
 inadmissible estimator, 375, 377
 increasing function, 64
 independence of random variables, 90, 93
 joint density factorization for, 91, 94
 implies covariance equals zero, 151
 of linear and quadratic forms in normally distributed random variables, 328
 of random vectors and functions of random vectors, 95
 independent events, 29, 32
 global independence of events, 90
 independent or explanatory variables, 429
 index set, 686
 indicator function, $I_A(x)$, 695
 induced likelihood function, 486
 induced probability space, 47
 inequalities, Cauchy-Schwarz, 149
 Chebyshev, 133–134
 Jensen's Inequality, 120, 125

Kolmogorov's inequality, 263
 Markov's inequality, 133
 instrumental variables estimator, 499
 integer programming, 469
 interval set notation, 687
 intervals, closed, 687
 half-open (or half-closed) intervals, 687
 open, 687
 inverse function from B to A , $f^{-1} : B \rightarrow A$, 693
 inverse image of y , $f^{-1}(y)$, 693
 inversion relationship between MGF and PDF, 144
 iso-density contours, 203
 isopower contour, 592

J

Jacobian matrix, 288, 340
 Jensen's Inequality, 120, 125
 joint density of the random sample, 302, 305
 when sampling from an exponential class distribution, 404
 joint cumulative distribution function, 73
 joint probability density function, 68

K

Khinchin's weak law of large numbers (WLLN), 259
 Kolmogorov's axioms of probability, 13
 inequality, 263
 SLLN, 264, 265
 Kolmogorov-Smirnov (K-S) test, 658
 Kronecker's Lemma, 265

L

Lagrange multiplier test, 616
 asymptotic distribution when $H_0 : R(\Theta) = r$ is true, 617
 asymptotic distribution when $H_a : R(\Theta) \neq r$ is true for local alternatives, 621

of $H_0 : R(\Theta_n) = r$ versus $H_a : R(\Theta_0) \neq r$, 616
 of $H_0 : \alpha = 1$ (exponential family) versus $H_a : \alpha \neq 1$ in gamma population distribution, 619
 least-squares estimator, 275, 437
 as a GMM estimator, 496
 as a maximum likelihood estimator, 468
 MVUE property of $(\hat{\beta}, \hat{S}^2)$ under normality, 462
 Lebesgue measure zero, 388
 Lehmann-Scheffé's completeness theorem, 421
 minimal sufficiency theorem, 395
 MVUE approach, 421
 Leibniz's rules for differentiation of integrals, 333
 level- α test, 532
 Liapounov CLT, 276
 for triangular arrays, 280
 Zolotarev's bound on approximation error, 278
 likelihood function, 464
 likelihood ratio test, 545
 limit of a real number sequence, 224
 of a real-valued matrix sequence, 226
 limiting CDF of $\{Y_n\}$, 234
 limiting density of $\{Y_n\}$, 236
 limiting distribution, 236
 relationship with asymptotic distribution, 239, 272
 limiting function of $\{f_n\}$ on D_0 , 230
 limiting moment-generating functions, 238
 Lindberg conditions, 274
 Lindberg CLT, 274
 Lindberg-Levy CLT, 270, 283
 Van Beeck's bound on approximation error, 273–274
 linear combinations of random variables, means and variance, 156
 of normal random variables, 206
 linear in parameters, 428
 local alternative hypotheses, 613
 location parameter family of PDFs, 652
 location-scale family of PDFs, 652

log-normal distribution, 216, 337, 358, 422, 425

M

m-dependence, 281
 maintained hypothesis, 529
 marginal
 cumulative distribution functions, 83
 cumulant-generating function, 147
 densities for $N(\mu, \Sigma)$, 207
 moment generating function, 147
 probability density, 78, 82
 Markov's inequality, 133
 mathematical induction, 11
 matrix A is smaller than matrix B , $A \preceq B$, 379
 matrix differentiation, 433, 701
 maximum likelihood estimator (MLE), 465
 as a GMM Estimator, 496
 asymptotic efficiency, 482
 asymptotic normality, 480
 attainment of the CRLB, 470
 consistency, 475
 for exponential distribution, 467
 for gamma distribution, 473
 for hypergeometric distribution, 469
 for normal distribution, 467
 for uniform distribution, 469
 in the GLM, 468
 invariance principle, 487
 MVUE property of unique MLEs, 471
 unique MLEs are functions of sufficient statistics, 470
 mean of a random variable, 132
 mean square error (scalar case), 373
 matrix, 374
 mean value theorem for integrals, 87
 median of X , 139
 memoryless property, geometric density 176
 exponential density, 191
 mesh, 114

method of moments (MOM)
 estimator, 491
 minimal sufficient statistic, 394
 minimum mean-square error
 criterion, 378
 minimum variance unbiased
 estimator (MVUE), 380
 necessary and sufficient
 conditions for, 406
 relationships between scalar and
 vector MVUEs, 406
 uniqueness of MVUEs, 408
 mixed discrete-continuous random
 variables, 59
 continuous density function
 component, 59
 discrete density function
 component, 59
 mixture distribution, 666
 mode of $f(x)$, 141
 model, statistical, 365
 moment generating function
 (MGF), 141, 146
 marginal, 147
 properties of, 143
 uniqueness theorem, 144, 147
 moments, 132
 about the mean (or central
 moment), 132
 about the origin, 132
 central moments as functions of
 moments about the origin, 137
 existence of moments, 138
 joint moment about the origin and
 mean, 148
 joint sample moment about the
 origin and mean, 321, 323
 moments about the origin as
 functions of central moments,
 138
 sample moments about the origin
 and mean, 316, 321
 monotone likelihood ratio in
 statistic $T = t(\mathbf{X})$, 551
 and the exponential class of
 densities, 551
 and UMP tests, 553
 multicollinearity, 457
 multinomial density, 174
 multiplication rule, 27, 28

N
 natural numbers, 222
 negative binomial density, 176
 Neyman-Pearson lemma, 540
 Neyman's factorization theorem,
 392
 noncentral F -distribution, 667
 t -distribution, 583, 628
 χ^2 -distribution, 666
 noncentrality parameter λ , 583
 nondecreasing function, 64
 nonparametric estimation, 364
 normal family of densities 197, 202
 nuisance parameters, 571
 null hypothesis, 527

O
 observable random variable, 307
 one-sided confidence interval or
 confidence bound, 641
 open interval, 687
 rectangle, 21
 operating characteristic function of
 a statistical test, 531
 order statistics, 351–352, 397
 order of magnitude in probability
 ($O_p(n^k)$ or $o_p(n^k)$), 248
 order of magnitude of a sequence,
 231
 of sums and products of
 sequences, 232
 outcome, of the experiment, 2
 of the random sample, 302, 305
 of the random variable, 45

P
 p -values, 535
 pairwise independent, 32
 parameter identifiability, 368
 parameter space, 169
 parameterization of a family of
 PDFs, 169
 parameters, 169
 parametric estimation, 363
 distribution-free case, 364
 distribution-specific case, 364
 parametric families of density
 functions, 169

parametric hypothesis testing, 528
 Pareto family of PDFs, 216, 591
 partition of a set, 10
 partitioned determinants, 209–210
 inversion, 209–210, 440
 Pascal distribution, 176
 piecewise invertible, 338
 Pittman drift, 613
 pivotal quantities, 650
 for location and/or scale
 parametric families of PDFs,
 652
 method of confidence-region
 construction, 650
 plim operator, 241
 properties, 245
 is an element-wise operator, 243
 applied to a matrix, 243
 point estimate, 371
 point estimator, 370
 Poisson density, 178
 Poisson process, 180–182
 mean rate of occurrence in, 182
 population, 298–299
 distribution, 300–302
 moments, 314
 posterior probabilities, 36
 power function of a statistical test,
 530
 ideal power function, 531
 prior information models, 463
 prior probabilities, 36
 probability, 8
 limit (See convergence in
 probability, and plim)
 measure, 13
 set function, 13
 space, 15
 of event A , 13
 probability integral transformation,
 348
 process, 297–299
 product notation, $\prod_{i=1}^n$, 32

Q
 quadratic formula for solving
 $ax^2 + bx + c = 0$, 161
 quantile of X , 140

R

R-squared, 434
 random number generator for standard normal PDF, 207
 for uniform PDF, 186–187, 350
 random sample, 305
 random sampling, from a population distribution, 301
 with replacement, 301
 without replacement, 303
 generated by a composite experiment, 305
 random variable, 44, 67
 random walk, 359
 randomized tests, 541, 555
 range of a random variable, $R[X]$, 45, 71
 range of X over the parameter space Ω , 393
 over H , 514
 Rao-Blackwell Theorem, 399, 400
 real line, 685
 real-valued function, 692
 point function, 694
 set function, 694
 vector function, 67
 rectangles in R^n , 21
 regression function, 211, 431
 curve, 129
 hyperplane, 211
 line, 211
 multivariate, 130
 of Y on X , 129
 rejection or critical region of statistical test, 515
 of test statistic, 526
 relation from A to B , $S : A \rightarrow B$, 690
 relative efficiency of estimators, 375, 377
 relative frequency probability, 6
 relatively more efficient, 375, 377
 reparameterization, of a probability density family, 485
 of exponential class densities via linear transformations, 579
 reparameterized, 170
 residual vector in the GLM, 430
 ridge regression estimator, 505
 runs, 663

S

sample, correlation, 324
 correlation matrix, 327, 328
 covariance, 321
 covariance matrix, 327
 distribution function, 308
 mean, 317, 321
 median, 353
 moments, 316, 321
 moments about the origin and mean, 314–316, 321
 points, 3
 space, 2
 standard deviation, 324
 variance, 319, 321
 sampling, densities, 332
 distributions, 332
 distribution of order statistics, 352–354
 error, 300
 scale parameter family of PDFs, 652
 scoring test (form of LM test), 617
 sequence of probability density functions, 224
 of random variables, 223
 sequences, 222
 adding, subtracting, and multiplying, 227
 asymptotic nonpositively correlated, 266
 bounded m -dependent sequences, 281
 bounded sequence of real numbers, 225
 combinations of, 227
 divergent, 224
 diverges to infinity, 225
 limit of a real number sequence, 224
 limit of a real-valued matrix sequence, 226
 order of magnitude of, 231–232
 sets, 679
 countable set, 681
 elements or members of, 679
 infinite set, 681
 multiple union and intersection notation, 686

mutually exclusive (or disjoint), 682
 operations and relationships, 682–684
 partition of, 10
 subset of, 682
 uncountably infinite, 681
 universal set, 682
 set-defining methods, 680
 exhaustive listing, 680
 mathematical rule, 680
 verbal rule, 680
 Shapiro-Wilks (SW) test of normality, 661
 sigma-algebra (σ -algebra), 21
 sigma-field (σ -field), 21
 significance level, 532
 simple hypothesis, 510
 simulating outcomes of random samples, 350
 simultaneous equations estimation, 459
 size of a statistical test, 524, 532
 size-of-set function, $N[A]$, 695
 skewed density, 136
 to the left or right, 136
 skewness, 136
 Slutsky's theorems, 248
 sociometrics, 460
 square root matrix (symmetric), V_n , 289
 standard deviation, of a random variable, 133
 of a sample, 324
 standard normal density and distribution, $N[0, 1]$, 198
 random-number generators for, 207
 random variable, 199
 standardizing a random variable, 199
 statistic, 307–308
 statistical hypothesis, 510
 inference, 297
 test, 515
 test defined by C_r , 516
 statistical model, 365
 stochastic process, 267, 299

strong law of large numbers (SLLN), 258, 263, 268
 Kolmogorov's, iid case, 264, 265
 non-iid case, 268
 strong mean square error (SMSE)
 superiority, 376
 subjective probability, 7
 consistency condition for, 18
 substitution theorem for
 conditional expectations, 127, 131
 sufficient statistics, 389
 complete, 400
 Lehmann-Scheffé's completeness theorem, 420
 Lehmann-Scheffé's minimal sufficiency theorem, 395
 minimal, 394
 Rao-Blackwell Theorem, 398, 400
 role in defining MVUEs, 420
 sufficiency and the MSE Criterion, 398
 sufficiency of invertible functions of sufficient statistics, 404
 support of a probability density function, 335, 541
 symmetric density about μ , 136

T

t probability density function, 341
 converges to standard normal,
 $T_v \xrightarrow{d} N(0, 1)$, 343
 Taylor series, expansion and remainder (Young's Form), 286
 representation of vector functions, 617
 test construction, heuristic principle of, 596
 test of a statistical hypothesis, 515
 most powerful, 534
 rejecting at the α -level of significance, 533
 rejecting using a size- α test, 533
 test statistic, 526
 testing functional forms of probability distributions, 655–663
 goodness-of-fit test (chisquare), 655

Kolmogorov-Smirnov (K-S) test, 658
 Lilliefors test for normality, 660
 normality of GLM error terms, 662
 Shapiro-Wilks (SW) test for normality, 661
 testing for a continuous uniform population distribution, 660
 discrete uniform population distribution, 673
 exponential population distribution, 657
 normal population distribution, 660
 tests of hypotheses, 510–674
 differences between two populations, means (equal variances), 669
 means (unequal variances–Behrens-Fisher problem), 669
 variances, 669
 equality of proportions, 673
 equality of two exponential population distributions, 670
 general linear model parameters:
 for σ^2 , asymptotic χ^2 and normal tests, 639–640
 for σ^2 , finite sample χ^2 -tests, 633–635
 $R(\beta) = r$, $R(\beta) \leq r$, or $R(\beta) \geq r$, scalar asymptotic normal tests, 636
 $R(\beta) = r$ or $R(\beta) = r$ when R has q rows, asymptotic χ^2 -tests, 636
 $R(\beta) = r$, $R\beta \leq r$, or $R\beta \geq r$ when R is $(1 \times k)$, t -tests, 628
 $R\beta = r$ when R is $q \times k$, F -tests, 626
 GLR test of $H_0 : R(\Theta_0) = r$ versus $H_a : R(\Theta_0) \neq r$, 610
 independence of bivariate normal random variables, 673
 in the multiparameter exponential class, 574
 LM Test of $H_0 : R(\Theta_0) = r$ versus $H_a : R(\Theta_0) \neq r$, 616

mean of a Bernoulli population distribution, $H_0 : p = p_0$ versus $H_a : p \neq p_0$, 559
 mean of a normal population distribution, 561, 566, 579, 586
 $H_0 : \mu = \mu_0$ versus $H_a : \mu \neq \mu_0$, 561, 579, 586
 $H_0 : \mu \leq \mu_0$ versus $H_a : \mu > \mu_0$, 579, 586
 $H_0 : \mu \in [\mu_1, \mu_2]$ versus $H_a : \mu \notin [\mu_1, \mu_2]$, 566
 mean of an exponential population distribution, 564, 568
 $H_0 : \theta = \theta_0$ versus $H_a : \theta \neq \theta_0$, 564
 $H_0 : \theta \in [\theta_1, \theta_2]$ versus $H_a : \theta \notin [\theta_1, \theta_2]$, 568
 paired comparisons of means, 671
 Wald Test of $H_0 : R(\Theta_0) = r$ versus $H_a : R(\Theta_0) \neq r$, 622–623
 there exists, \exists , 691
 triangular array of random variables, 279
 true PDF of X , 364
 true value of Θ , 363
 of $q(\Theta)$, 364
 truncation function, $\text{trunc}(x)$, 62
 type I and type II error, 516
 controlling, 519
 protection against type I error, 523

U

unbiased, confidence region, 648
 estimator, 378
 statistical test, 533
 uncorrelated, 156
 uniform density function,
 continuous, 186
 discrete, 170
 uniform random-number generators, 186–187, 350
 uniformly minimum variance unbiased estimator (UMVUE), 380
 uniformly most accurate (UMA) confidence region, 648

- uniformly most accurate unbiased [UMAU] confidence regions, 649
- uniformly most powerful (UMP) test, 534
- GLR test of $H_0 : \Theta = \Theta_0$ versus $H_a : \Theta \in \Omega_a$, 604
- GLR test of $H_0 : \Theta \in \Omega_0$ versus one-sided $H_a : \Theta \in \Omega_a$, 606
- in terms of the $t(X)$ of a monotone likelihood ratio, 554
- nonexistence of a UMP test, 548, 556
- of $H_0 : \Theta = \Theta_0$ versus $H_a : \Theta \in \Omega_a$ using Neyman-Pearson lemma, 547
- unbiasedness of UMP tests, 546, 550, 554
- uniformly most powerful unbiased (UMPU) test, 556
- in multiple parameter cases via conditioning, 571
- of $H_0 : \sigma^2 \leq \sigma_0^2$ versus $H_a : \sigma^2 > \sigma_0^2$ when sampling from a normal population distribution, 579
- of $H_0 : \mu \leq \mu_0$ versus $H_a : \mu > \mu_0$, normal population distribution, 586
- of $H_0 : \mu = \mu_0$ versus $H_a : \mu \neq \mu_0$, normal population distribution, 586
- of $H_0 : \Theta = \Theta_0$ versus $H_a : \Theta \neq \Theta_0$ when test statistics have symmetric PDFs, 569
- two-sided tests for exponential class densities, 558
- unimodal, 141
- V**
- variance of a random variable, 133
- sample variance, 319, 322
- W**
- Wald test, 622–623
- asymptotic distribution when $H_0 : R(\Theta) = r$ is true, 623
- asymptotic distribution when $H_a : R(\Theta) \neq r$ is true for local alternatives, 624
- based on a GMM estimator, 624
- Wald-Wolfowitz (WW) runs test of iid, 663
- weak mean-square error (WMSE), 377
- weak law of large numbers (WLLN), 258
- Khinchin's, iid case, 259
- necessary and sufficient conditions for, 261
- non-iid case, 262
- wear-out effects, 191
- Weierstrass's approximation theorem, 429
- work-hardening effects, 191