Yuan Shih Chow
Henry Teicher

Probability Theory

Independence
Interchangeability
Martingales

Yuan Shih Chow
Henry Teicher

Probability Theory

Independence
Interchangeability
Martingales

Springer-Verlag
New York Heidelberg Berlin

Yuan Shih Chow

Department of Mathematics and Statistics
Columbia University
New York, New York 10027
USA

Henry Teicher

Department of Statistics
Rutgers University
New Brunswick, New Jersey 08903
USA

Library of Congress Cataloging in Publication Data

Chow, Yuan-shih, 1924-
Probability theory: independence, interchangeability, Martingales.

Includes bibliographies and index.
1. Probabilities. 2. Martingales (Mathematics)
I. Teicher, Henry, joint author. II. Title.
QA273.C573 519.2 78-21693
ISBN 0-387-90331-3

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-90331-3 Springer-Verlag New York
ISBN 3-540-90331-3 Springer-Verlag Heidelberg Berlin

Preface

Probability theory is a branch of mathematics dealing with chance phenomena and has clearly discernible links with the real world. The origins of the subject, generally attributed to investigations by the renowned french mathematician Fermat of problems posed by a gambling contemporary to Pascal, have been pushed back a century earlier to the italian mathematicians Cardano and Tartaglia about 1570 (Ore, 1953). Results as significant as the Bernoulli weak law of large numbers appeared as early as 1713, although its counterpart, the Borel strong law of large numbers, did not emerge until 1909. Central limit theorems and conditional probabilities were already being investigated in the eighteenth century, but the first serious attempts to grapple with the logical foundations of probability seem to be Keynes (1921), von Mises (1928; 1931), and Kolmogorov (1933).

An axiomatic mold and measure-theoretic framework for probability theory was furnished by Kolmogorov. In this so-called objective or measure-theoretic approach, definitions and axioms are so chosen that the empirical realization of an event is the outcome of a not completely determined physical experiment—an experiment which is at least conceptually capable of indefinite repetition (this notion is due to von Mises). The concrete or intuitive counterpart of the probability of an event is a long run or limiting frequency of the corresponding outcome.

In contradistinction to the objective approach—where typical realizations of events might be: a coin falls heads, more than 50 cars reach a busy intersection during a specified period, a continuously burning light bulb fails within 1000 hours—the subjective approach to probability advocated by Keynes is designed to encompass realizations such as: it will rain tomorrow, life exists on the planet Saturn, the Iliad and the Odyssey were written by the same author—despite the fact that the experiments in question are clearly unrepeatable. Here the empirical counterpart of probability is degree or intensity of belief.

It is tempting to try to define probability as a limit of frequencies (as advocated by von Mises) rather than as a real number between zero and one satisfying certain postulates (as in the objective approach). Unfortunately, incorporation of repeatability as a postulate (von Mises' "randomness axiom") complicates matters while simultaneously circumscribing the notion of an event. Thus, the probability of the occurrence infinitely often of some particular event in an infinite sequence of repetitions of an experiment—which is of considerable interest in the Kolmogorov schema—is proscribed in (the 1964 rendition of) the von Mises approach (1931). Possibly for these reasons, the frequency approach appears to have lost out to the measure-theoretic. It should be pointed out, however, that justification of the measure-theoretic approach via the Borel strong law of large numbers is circular in that the convergence of the observed frequency of an event to its theoretically defined probability (as the number of repetitions increases) is not pointwise but can only be defined in terms of the concept being justified, viz., probability. If, however, one is willing to ascribe an intuitive meaning to the notion of probability one (hence also, probability zero) then the probability p of any intermediate value can be interpreted in this fashion.

A number of axiomatizations for subjective probability have appeared since Keynes with no single approach dominating. Perhaps the greatest influence of subjective probability is outside the realm of probability theory proper and rather in the recent emergence of the Bayesian school of statistics.

The concern of this book is with the measure-theoretic foundations of probability theory and (a portion of) the body of laws and theorems that emerge thereform. In the 45 years since the appearance of von Mises' and Kolmogorov's works on the foundations of probability, the theory itself has expanded at an explosive pace. Despite this burgeoning, or perhaps because of the very extent thereof, only the topics of independence, interchangeability, and martingales will be treated here. Thus, such important concepts as Markov and stationary processes will not even be defined, although the special cases of sums of independent random variables and interchangeable random variables will be dealt with extensively. Likewise, continuous parameter stochastic processes, although alluded to, will not be discussed. Indeed, the time seems propitious for the appearance of a book devoted solely to such processes and presupposing familiarity with a significant portion of the material contained here.

Particular emphasis is placed in this book on stopping times—on the one one hand, as tools in proving theorems, and on the other, as objects of interest in themselves. Apropos of the latter, randomly stopped sums, optimal stopping problems, and limit distributions of sequences of stopping rules (i.e., finite stopping times) are of special interest. Wald's equation and its second-moment analogue, in turn, show the usefulness of such stopped sums in renewal theory and elsewhere in probability. Martingales provide a natural vehicle for stopping times, but a formal treatment of the latter cannot await development of the former. Thus, stopping times and, in particular, a sequence of copies of a fixed stopping rule appear as early as Chapter 5, thereby facilitating discussion of the limiting behavior of random walks.

Many of the proofs given and a few of the results are new. Occasionally, a classical notion is looked at through new lenses (e.g., reformulation of the Lindeberg condition). Examples, sprinkled throughout, are used in various guises; to extend theory, to illustrate a theorem that has just appeared, to obtain a classical result from one recently proven.

A novel feature is the attempt to intertwine measure and probability rather than, as is customary, set up between them a sharp demarcation. It is surprising how much probability can be developed (Chapters 2, 3) without even a mention of integration. A number of topics treated later in generality are foreshadowed in the very tractable binomial case of Chapter 2.

This book is intended to serve as a graduate text in probability theory. No knowledge of measure or probability is presupposed, although it is recognized that most students will have been exposed to at least an elementary treatment of the latter. The former is confined for the most part to Chapters 1, 4, 6, with convergence appearing in Section 3.3 (i.e., Section 3 of Chapter 3).[1] Readers familiar with measure theory can plunge into Chapter 5 after reading Section 3.2 and portions of Sections 3.1, 3.3, 4.2, 4.3. In any case, Chapter 2 and also Section 3.4 can be omitted without affecting subsequent developments.

Martingales are introduced in Section 7.4, where the upward case is treated and then developed more generally in Chapter 11. Interchangeable random variables are discussed primarily in Sections 7.3 and 9.2. Apropos of terminology, "interchangeable" is far more indicative of the underlying property than the current "exchangeable," which seems to be a too literal rendition of the french word "échangeable."

A one-year course presupposing measure theory can be built around Chapters 5, 7, 8, 9, 10, 11, 12.

Our warm thanks and appreciation go to Mary Daughaday and Beatrice Williams for their expert typing of the manuscript.

References

J. M. Keynes, *A Treatise on Probability*, 1921; MacMillan, London, 1943.

A. Kolmogorov, *Foundations of the Theory of Probability*, 1933; Chelsea, New York, 1950.

R. von Mises, *Probability, Statistics and Truth*, 1928; Wm. Hodge, London, 1939.

R. von Mises, *Mathematical Theory of Probability and Statistics*, 1931 (H. Geiringer, editor), Academic Press, N.Y., 1964.

O. Ore, "Appendix," *Cardano, The Gambling Scholar*, Princeton University Press, 1953; Holt, New York, 1961.

I. Todhunter, *A History of the Mathematical Theory of Probability*, 1865; Chelsea, New York, 1949.

[1] In the same notational vein, Theorem 3.4.2 signifies Theorem 2 of Section 4 of Chapter 3.

List of Abbreviations

r.v	random variable
r v s	random variables
d f	distribution function
c f.	characteristic function
p.d.f.	probability density function
u.i.	uniform integrability
i.o.	infinitely often
a.c.	almost certainly
a.s.	almost surely
a.e.	almost everywhere
i.d.	infinitely divisible
i.i.d.	independent, identically distributed
iff	if and only if
CLT	Central Limit Theorem
WLLN	Weak Law of Large Numbers
SLLN	Strong Law of Large Numbers
LIL	Law of the Iterated Logarithm

List of Symbols and Conventions

$\sigma(\mathscr{G})$	σ-algebra generated by the class $\mathscr{G}$
$\sigma(X)$	σ-algebra generated by the random variable X
$\mathrm{E}\,X$	expectation of the random variable X
$\int X$	abbreviated form of the integral $\int X\,dP$
$\mathrm{E}^p\,X$	abbreviated form of $(\mathrm{E}\,X)^p$
$\|X\|_p$	p-norm of X, that is, $(\mathrm{E}\,\|X\|^p)^{1/p}$
$\mathrm{C}(F)$	continuity set of the function F
$\xrightarrow{\text{a.c. or a.s. or a.e.}}$	convergence almost certainly or almost surely or almost everywhere
$\xrightarrow{P \text{ or } d \text{ or } \mu}$	convergence in probability or in distribution or in μ-measure
$\xrightarrow{\mathscr{L}_p}$	convergence in mean of order p
$\xrightarrow{\text{w or c}}$	weak or complete convergence
$\mathscr{B}^n$ or $\mathscr{B}^\infty$	class of n-dimensional or infinite-dimensional Borel sets
$\mathscr{R}\{\ \}$	real part of
$\mathscr{I}\{\ \}$	imaginary part of
$\wedge$	minimum of
$\vee$	maximum of
$a \le \underline{\overline{\lim}}\, Y_n \le b$	simultaneous statement that $a \le \underline{\lim}_{n\to\infty} Y_n \le \overline{\lim}_{n\to\infty} Y_n \le b$
$Z \le z_2$, $\ge z_1$	simultaneous statement that $Z \le z_2$ and $Z \ge z_1$

Contents

Classes of Sets, Measures, and Probability Spaces 1

1.1 Sets and Set Operations

A **set** in the words of Georg Cantor, the founder of modern set theory, is a collection into a whole of definite, well-distinguished objects of our perception or thought, The objects are called elements and the set is the aggregate of these elements. It is very convenient to extend this notion and also envisage a set devoid of elements, a so-called empty set, and this will be denoted by $\varnothing$. Each element of a set appears only once therein and its order of appearance within the set is irrelevant. A set whose elements are themselves sets will be called a class.

Examples of sets are (i) the set of positive integers denoted by either $\{1, 2, \ldots\}$ or $\{\omega: \omega \text{ is a positive integer}\}$ and (ii) the closed interval with end points a and b benoted by either $\{\omega: a \leq \omega \leq b\}$ or $[a, b]$. Analogously, the open interval with end points a and b is denoted by $\{\omega: a < \omega < b\}$ or (a, b), while $(a, b]$ and $[a, b)$ are designations for $\{\omega: a < \omega \leq b\}$ and $\{\omega: a \leq \omega < b\}$ respectively.

The statement that $\omega \in A$ means that ω is an element of the set A and analogously the assertion $\omega \notin A$ means that ω is not an element of the set A or alternatively that ω does not belong to A. If A and B are sets and every element of A is likewise an element of B, this situation is depicted by writing $A \subset B$ or $B \supset A$, and in such a case the set A is said to be a subset of B or contained in B. If both $A \subset B$ and $B \subset A$, then A and B contain exactly the same elements and are said to be equal, denoted by $A = B$. Note that for every set A, $\varnothing \subset A \subset A$.

A set A is termed **countable** if there exists a one-to-one correspondence between (the elements of) A and (those of) some subset B of the set of all positive integers. If, in this correspondence, $B = \{1, 2, \ldots, n\}$, then A is called a **finite** set (with n elements). It is natural to consider $\varnothing$ as a finite set (with zero elements). A set A which is not countable is called **uncountable** or **nondenumerable**.

If A and B are two sets, the difference $A - B$ is the set of all elements of A which do not belong to B; the intersection $A \cap B$ or $A \cdot B$ or simply AB is the set of all elements belonging to both A and B; the union $A \cup B$ is the set of all elements belonging to either A or B (or both); and the symmetric difference $A \,\Delta\, B$ is the set of all elements which belong to A or B but not both. Note that

$$A \cup A = A, \qquad A \cap A = A, \qquad A - A = \varnothing, \qquad A - B = A - (AB) \subset A,$$
$$A \cup B = B \cup A \supset A \supset AB = BA, \qquad A \,\Delta\, B = (A - B) \cup (B - A).$$

Union, intersection, difference, and symmetric difference are termed **set operations**.

If A, B, C are sets and several set operations are indicated, it is, strictly speaking, necessary to indicate via parentheses which operations are to be performed first. However, such specification is frequently unnecessary. For instance, $(A \cup B) \cup C = A \cup (B \cup C)$ and so this double union is independent of order and may be designated simply by $A \cup B \cup C$. Analogously,

$$(AB)C = A(BC) = ABC, \qquad (A \,\Delta\, B) \,\Delta\, C = A \,\Delta\, (B \,\Delta\, C) = A \,\Delta\, B \,\Delta\, C,$$
$$A(B \cup C) = AB \cup AC, \qquad A(B \,\Delta\, C) = AB \,\Delta\, AC.$$

If Λ is a nonempty set whose elements λ may be envisaged as tags or labels, $\{A_\lambda : \lambda \in \Lambda\}$ is a nonempty class of sets. The intersection $\bigcap_{\lambda \in \Lambda} A_\lambda$ (resp. union $\bigcup_{\lambda \in \Lambda} A_\lambda$) is defined to be the set of all elements which belong to A_λ for all $\lambda \in \Lambda$ (resp. for some $\lambda \in \Lambda$). Apropos of order of carrying out set operations, if $*$ denotes any one of $\cup, \cap, -, \Delta$, for any set A it follows from the definitions that

$$\bigcup_{\lambda \in \Lambda} A_\lambda * A = \left(\bigcup_{\lambda \in \Lambda} A_\lambda\right) * A, \qquad A * \bigcup_{\lambda \in \Lambda} A_\lambda = A * \left(\bigcup_{\lambda \in \Lambda} A_\lambda\right),$$
$$\bigcap_{\lambda \in \Lambda} A_\lambda * A = \left(\bigcap_{\lambda \in \Lambda} A_\lambda\right) * A, \qquad A * \bigcap_{\lambda \in \Lambda} A_\lambda = A * \left(\bigcap_{\lambda \in \Lambda} A_\lambda\right).$$

Then

$$A - \bigcup_{\lambda \in \Lambda} A_\lambda = \bigcap_{\lambda \in \Lambda} (A - A_\lambda), \qquad A - \bigcap_{\lambda \in \Lambda} A_\lambda = \bigcup_{\lambda \in \Lambda} (A - A_\lambda).$$

For any sequence $\{A_n, n \geq 1\}$ of sets, define

$$\underline{\lim}_{n \to \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k, \qquad \overline{\lim}_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

and note that, employing the abbreviation i.o. to designate "infinitely often,"

$$\overline{\lim}_{n \to \infty} A_n = \{\omega : \omega \in A_n \text{ for infinitely many } n\} = \{\omega : \omega \in A_n, \text{ i.o.}\}$$
$$\underline{\lim}_{n \to \infty} A_n = \{\omega : \omega \in A_n \text{ for all but a finite number of indices } n\}. \tag{1}$$

To prove, for example, the first relation, let $A = \{\omega: \omega \in A_n, \text{ i.o.}\}$. Then $\omega \in A$ iff for every positive integer m, there exists $n \geq m$ such that $\omega \in A_n$, that is, iff for every positive integer m, $\omega \in \bigcup_{n=m}^{\infty} A_n$ i.e., iff $\omega \in \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$.

In view of (1), $\underline{\lim}\, A_n \subset \overline{\lim}\, A_n$, but these two sets need not be equal (Exercise 3). If $\underline{\lim}\, A_n = \overline{\lim}\, A_n = A$ (say), A is called the limit of the sequence A_n; this situation is depicted by writing $\lim A_n = A$ or $A_n \to A$. If $A_1 \subset A_2 \subset \cdots$ (resp. $A_1 \supset A_2 \supset \cdots$) the sequence A_n is said to be **increasing** (resp. **decreasing**). In either case, $\{A_n, n \geq 1\}$ is called **monotone**.

Palpably, for every monotone sequence A_n, $\lim_{n\to\infty} A_n$ exists; in fact, if $\{A_n\}$ is increasing, $\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$, while if $\{A_n\}$ is decreasing, $\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$. Consequently, for any sequence of sets A_n,

$$\overline{\lim_{n\to\infty}}\, A_n = \lim_{n\to\infty} \bigcup_{k=n}^{\infty} A_k,$$

$$\underline{\lim_{n\to\infty}}\, A_n = \lim_{n\to\infty} \bigcap_{k=n}^{\infty} A_k.$$

Exercises 1.1

1. Prove (i) if A_n is countable, $n \geq 1$, so is $\bigcup_{n=1}^{\infty} A_n$; (ii) if A is uncountable and $B \supset A$, then B is uncountable.
2. Show that $\bigcup_{n=1}^{\infty} [0, n/(n+1)) = [0, 1)$, $\bigcap_{n=1}^{\infty} (0, 1/n) = \varnothing$.
3. Prove that $\underline{\lim}_{n\to\infty} A_n \subset \overline{\lim}_{n\to\infty} A_n$. Specify $\overline{\lim}\, A_n$ and $\underline{\lim}\, A_n$ when $A_{2j} = B$, $A_{2j-1} = C, j = 1, 2, \ldots$.
4. Verify that $\bigcup_{n=1}^{\infty} A_n = \lim_{n\to\infty} \bigcup_1^n A_j$ and $\bigcap_{n=1}^{\infty} A_n = \lim_{n\to\infty} \bigcap_{j=1}^{n} A_j$. Moreover, if $\{A_n, n \geq 1\}$ is a sequence of disjoint sets, i.e., $A_i A_j = \varnothing, i \neq j$, then
$$\lim_{n\to\infty} \bigcup_{j=n}^{\infty} A_j = \varnothing.$$
5. Prove that $\overline{\lim}_n (A_n \cup B_n) = \overline{\lim}_n A_n \cup \overline{\lim}_n B_n$ and $\underline{\lim}_n A_n \cdot B_n = \underline{\lim}_n A_n \cdot \underline{\lim}_n B_n$. Moreover, $\lim A_n = A$ and $\lim B_n = B$ imply $\lim_n (A_n \cup B_n) = A \cup B$ and
$$\lim A_n B_n = AB.$$
6. Demonstrate that if B is a countable set and $B_n = \{(b_1, \ldots, b_n): b_i \in B \text{ for } 1 \leq i \leq n\}$, then B_n is countable, $n \geq 1$.
7. Prove that the set S consisting of all infinite sequences with entries 0 or 1 is nondenumerable and conclude that the set of real numbers in $[0, 1]$ or any nondegenerate interval is nondenumerable. *Hint*: If S were countable, i.e., $S = \{s_n, n \geq 1\}$ where $s_n = (x_{n1}, x_{n2}, \ldots)$, then $(1 - x_{11}, 1 - x_{22}, \ldots, 1 - x_{nn}, \ldots)$ would be an infinite sequence of zeros and ones not in S.
8. If a_n is a sequence of real numbers, $0 \leq a_n \leq \infty$, prove that
$$\bigcup_{n=1}^{\infty} [0, a_n) = \left[0, \sup_{n\geq 1} a_n\right), \qquad \bigcup_{n=1}^{\infty} \left[0, \left(\frac{n+1}{n}\right)^n\right] \neq \left[0, \sup_{n\geq 1} \left(\frac{n+1}{n}\right)^n\right].$$
9. For any sequence of sets $\{A_n, n \geq 1\}$, define $B = A_1$, $B_{n+1} = B_n \,\Delta\, A_{n+1}$, $n \geq 1$. Prove that $\lim_n B_n$ exists iff $\lim A_n$ exists and is empty.

1.2 Spaces and Indicators

A **space** Ω is an arbitrary, nonempty set and is usually postulated as a reference or point of departure for further discussion and investigation. Its elements are referred to as points (of the space) and will be denoted generically by ω. Thus $\Omega = \{\omega: \omega \in \Omega\}$.

For any reference space Ω, the **complement** A^c of a subset A of Ω is defined by $A^c = \Omega - A$ and the **indicator** I_A of $A \subset \Omega$ is a function defined on Ω by

$$I_A(\omega) = 1 \quad \text{for } \omega \in A, \qquad I_A(\omega) = 0 \quad \text{for } \omega \in A^c.$$

Similarly, for any real function f on Ω and real constants a, b, $I_{[a \le f \le b]}$ signifies the indicator of the set $\{\omega: a \le f(\omega) \le b\}$.

For any subsets A, B of Ω

$$A \subset B \quad \text{iff } A^c \supset B^c,$$

$$(A^c)^c = A, \qquad A \cup A^c = \Omega, \qquad A \cdot A^c = \varnothing, \qquad I_{A^c} = 1 - I_A,$$

$$A - B = AB^c, \qquad I_A \le I_B \quad \text{iff } A \subset B, \qquad I_{A \cup B} \le I_A + I_B,$$

with the last inequality becoming an equality for all ω iff $AB = \varnothing$. Let Λ be an arbitrary set and $\{A_\lambda, \lambda \in \Lambda\}$ a class of subsets of Ω. It is convenient to adopt the conventions

$$\bigcup_{\lambda \in \varnothing} A_\lambda = \varnothing, \qquad \bigcap_{\lambda \in \varnothing} A_\lambda = \Omega.$$

Moreover,

$$\left(\bigcup_{\lambda \in \Lambda} A_\lambda\right)^c = \bigcap_{\lambda \in \Lambda} A_\lambda^c, \qquad \left(\bigcap_{\lambda \in \Lambda} A_\lambda\right)^c = \bigcup_{\lambda \in \Lambda} A_\lambda^c,$$

$$I_{\bigcap_{\lambda \in \Lambda} A_\lambda} = \inf_{\lambda \in \Lambda} I_{A_\lambda}, \qquad I_{\bigcup_{\lambda \in \Lambda} A_\lambda} = \sup_{\lambda \in \Lambda} I_{A_\lambda}.$$

If $A_\lambda \cdot A_{\lambda'} = \varnothing$ for $\lambda, \lambda' \in \Lambda$ and $\lambda \ne \lambda'$, the sets A_λ are called **disjoint**. A class of disjoint sets will be referred to as a **disjoint class**.

If $\{A_n, n \ge 1\}$ is a sequence of subsets of Ω, then $\{I_{A_n}, n \ge 1\}$ is a sequence of functions on Ω with values 0 or 1 and

$$I_{\underline{\lim}\, A_n} = \underline{\lim}_{n \to \infty} I_{A_n}, \qquad I_{\overline{\lim}\, A_n} = \overline{\lim}_{n \to \infty} I_{A_n}.$$

Moreover,

$$I_{\bigcup_1^\infty A_n} \le \sum_1^\infty I_{A_n}. \tag{1}$$

Equality holds in (1) iff $\{A_n, n \geq 1\}$ is a disjoint class. The following identity (2) is a refinement of the finite counterpart of (1): For $A_i \subset \Omega$, $1 \leq i \leq n$, set

$$s_1 = \sum_1^n I_{A_j}, \qquad s_2 = \sum_{1 \leq j_1 < j_2 \leq n} I_{A_{j_1}A_{j_2}}, \ldots,$$

$$s_n = \sum_{1 \leq j_1 < \cdots < j_n \leq n} I_{A_{j_1}A_{j_2}\cdots A_{j_n}} = I_{A_1A_2\cdots A_n}.$$

Then

$$I_{\cup_1^n A_j} = s_1 - s_2 + s_3 - \cdots + (-1)^{n-1}s_n. \tag{2}$$

In proof of (2), if for some $\omega \in \Omega$, $I_{\cup_1^n A_j}(\omega) = 0$, clearly $s_k(\omega) = 0$, $1 \leq k \leq n$, whence (2) obtains. On the other hand, if $I_{\cup_1^n A_j}(\omega) = 1$, then $\omega \in A_j$ for at least one j, $1 \leq j \leq n$. Suppose that ω belongs to exactly m of the sets $A_1, \ldots, A_n$. Then $s_1(\omega) = m$, $s_2(\omega) = \binom{m}{2}, \ldots, s_m(\omega) = 1$, $s_{m+1}(\omega) = \cdots = s_n(\omega) = 0$, whence

$$s_1 - s_2 + \cdots + (-1)^{n-1}s_n = m - \binom{m}{2} + \cdots + (-1)^{m-1}\binom{m}{m} = 1 = I_{\cup_1^n A_j}.$$

Exercises 1.2

1. Verify that

$$A \,\Delta\, B = A^c \,\Delta\, B^c, \qquad C = A \,\Delta\, B \quad \text{iff } A = B \,\Delta\, C,$$

$$\bigcup_1^\infty A_n \,\Delta\, \bigcup_1^\infty B_n \subset \bigcup_1^\infty (A_n \,\Delta\, B_n),$$

$$\bigcap_1^\infty A_n \,\Delta\, \bigcap_1^\infty B_n \subset \bigcap_1^\infty (A_n \,\Delta\, B_n).$$

2. Prove that $(\underline{\lim}_{n\to\infty} A_n)^c = \overline{\lim}_{n\to\infty} A_n^c$ and $(\overline{\lim}_{n\to\infty} A_n)^c = \underline{\lim}_{n\to\infty} A_n^c$.
3. Prove that $I_{\overline{\lim}\, A_n} = \overline{\lim}\, I_{A_n}$ and that $I_{\lim A_n} = \lim I_{A_n}$ whenever either side exists.
4. If $A_n \subset \Omega$, $n \geq 1$, show that

$$I_{\cap_{n=1}^\infty A_n} = \max_{n\geq 1} I_{A_n}, \qquad I_{\cap_{n=1}^\infty A_n} = \min_{n \geq 1} I_{A_n}.$$

5. If f is a real function on Ω, then $f^2 = f$ iff f is an indicator of some subset of Ω.
6. Apropos of (2), prove that if B_m is the set of points belonging to exactly $m(1 \leq m \leq n)$ of $A_1, \ldots, A_n$, then

$$I_{B_m} = s_m - \binom{m+1}{m}s_{m+1} + \binom{m+2}{m}s_{m+2} - \cdots + (-1)^{n-m}\binom{n}{m}s_n. \tag{3}$$

7. If $\{f_n, n \geq 0\}$ is a sequence of real functions on Ω with $f_n \uparrow f_0$ and $A_n = \{\omega: f_n(\omega) > c\}$, then $A_n \subset A_{n+1}$ and $\lim A_n = A_0$.

8. If $\{f_n, n \geq 0\}$ is a sequence of real functions with $f_n \uparrow f_0$ and $g_n = f_n I_{[a \leq f_n \leq b]}$ for some real constant $a < b$, then $\{g_n, n \geq 1\}$ is not necessarily increasing. However, if for $n \geq 0$
$$f'_n = f_n I_{[a \leq f_n \leq b]} + a I_{[f_n < a]} + b I_{[f_n > b]},$$
then $f'_n \uparrow f'_0$.

9. If f_1 and f_2 are real functions on Ω, prove that for all real x and rational r
$$\{\omega: f_1(\omega) + f_2(\omega) < x\} = \bigcup_{\text{all } r} \{\omega: f_1(\omega) < r\} \cdot \{\omega: f_2(\omega) < x - r\}.$$

1.3 σ-Algebras, Measurable Spaces, and Product Spaces

Let Ω be a space.

Definition. A nonempty class $\mathscr{A}$ of subsets of Ω is an **algebra** if

i. $A^c \in \mathscr{A}$ whenever $A \in \mathscr{A}$,
ii. $A_1 \cup A_2 \in \mathscr{A}$ whenever $A_j \in \mathscr{A}, j = 1, 2$.

Moreover, $\mathscr{A}$ is a **σ-algebra** if, in addition,

iii. $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$ whenever $A_n \in \mathscr{A}, n \geq 1$.

Evidently, (ii) implies that for every positive integer n, $\bigcup_1^n A_j \in \mathscr{A}$ whenever $A_j \in \mathscr{A}$, $1 \leq j \leq n$, while both (i) and (ii) entail $A_1 A_2 \in \mathscr{A}$, if $A_j \in \mathscr{A}$, $j = 1, 2$,; also, since $\mathscr{A}$ is nonempty, $\Omega \in \mathscr{A}$, $\varnothing \in \mathscr{A}$. Clearly, (iii) implies (ii) by taking $A_n = A_2$, $n \geq 2$. Note that a σ-algebra is closed under countable intersections.

Definition. A nonempty class $\mathscr{A}$ of subsets of Ω is a **monotone class** if $\lim A_n \in \mathscr{A}$ for every monotone sequence $A_n \in \mathscr{A}$, $n \geq 1$.

Obviously, a σ-algebra is a monotone class. Conversely, a monotone algebra $\mathscr{A}$ (i.e., a monotone class which is simultaneously an algebra) is a σ-algebra. For if $A_n \in \mathscr{A}$, $n \geq 1$, then $B_n = \bigcup_{j=1}^n A_j \in \mathscr{A}$, $n \geq 1$, whence $\bigcup_{j=1}^{\infty} A_j = \lim_n B_n \in \mathscr{A}$.

Let S_Ω be the class of all subsets of Ω and $T_\Omega = \{\varnothing, \Omega\}$. Then S_Ω and T_Ω are σ-algebras and for any σ-algebra U_Ω of subsets of Ω, $T_\Omega \subset U_\Omega \subset S_\Omega$.

Definition. The **minimal algebra** $\mathscr{E}'$ (resp. σ-algebra, monotone class) containing a nonempty class $\mathscr{E}$ of subsets of Ω, is an algebra (resp. σ-algebra, monotone class) such that

i. $\mathscr{E}' \supset \mathscr{E}$,