CONCEPTS OF PROBABILITY THEORY

SECOND REVISED EDITION

PAUL E. PFEIFFER

CONCEPTS OF PROBABILITY THEORY

PAUL E. PFEIFFER

Department of Mathematical Sciences
Rice University

Second Revised Edition

DOVER PUBLICATIONS, INC.
NEW YORK

to Mamma and Daddy

Who contributed to this book
by teaching me that hard work can be
satisfying if the task is worthwhile.

Bibliographical Note

This Dover edition, first published in 1978, is an unabridged republication of the work originally published by McGraw-Hill Book Company, New York, in 1965. This Dover edition has been extensively revised and updated by the author.

Library of Congress Catalog Card Number: 78-53757

International Standard Book Number
ISBN 13: 978-0-486-63677-1
ISBN 10: 0-486-63677-1

Manufactured in the United States by Courier Corporation
63677107
www.doverpublications.com

Preface

My purpose in writing this book is to provide for students of science, engineering, and mathematics a course at the junior or senior level which lays a firm theoretical foundation for the application of probability theory to physical and other real-world problems. Mastery of the material presented should provide an excellent background for further study in such rapidly developing areas as statistical decision theory, reliability theory, dynamic programming, statistical game theory, coding and information theory, communication and control in the presence of noise, etc., as well as in classical sampling statistics.

My teaching experience has shown that students of engineering and science can master the essential features of a precise mathematical model in a way that clarifies their thinking and extends their ability to make significant applications of the theory. While the ultimate aim of the development in this book is the application of probability theory to problems of practical import and interest, the central task undertaken is an exposition of the basic concepts of probability theory, including a substantial introduction to the idea of a random process. Rather than provide a treatise on applications and techniques, I have attempted to provide a clear development of the fundamental concepts and theoretical perspectives which guide the formulation of problems and the discovery of methods of solution. Considerable attention is given to the task of translating real-world problems into the precise concepts of the model, so that the problem is stated unambiguously and may be attacked with all the resources provided by the mathematical theory as well as physical insight.

The rich theory of probability may be constructed on the essentially simple conceptual framework of that mathematical model generally known as the Kolmogorov model. The central features of this model may be grasped with the aid of certain graphical, mechanical, and notational representations which facilitate the formulation and visualization of concepts and relationships. Highly sophisticated techniques are seen as the means of performing conceptually simple tasks. The whole theory is formulated in a way that makes contact with both the literature of applications and the literature on pure mathematics. At many points I have borrowed specifically and explicitly from one or the other; the resulting treatment points the way to extend such use of the vast reservoir of knowledge in this important and rapidly growing field.

In introducing the basic model, I have appealed to the notion of probability as an idealization of the concept of the relative frequency of occurrence of an event in a large number of repeated trials. In most places, the primary interpretation of probability has been in terms of this familiar concept. I have also made considerable use of the idea that probability indicates the uncertainty regarding the outcome of a trial before the result is known. Various thinkers are currently advocating other approaches to formulating the mathematical model, and hence to interpreting its features. These alternative ways of developing and interpreting the model do not alter its character or the strategies and techniques for dealing with it. They do serve to increase confidence in the usefulness and “naturalness” of the model and to point to the desirability of achieving a mastery of the theory based upon it.

The background assumed in the book is provided in the freshman and sophomore mathematics courses in many universities. A knowledge of limits, differentiation, and integration is essential. Some acquaintance with the rudiments of set theory is assumed in the text, but an Appendix provides a brief treatment of the necessary material to aid the student who does not have the required background. Many students from the high schools offering instruction in the so-called *new mathematics* will be familiar with most of the material on sets before entering the university. Although some applications are made to physical problems, very little technical background is needed to understand these applications. The book should be suitable for a course offered in an engineering or science department, or for a course offered by a department of mathematics for students of engineering or science. The practicing engineer or scientist whose formal education did not provide a satisfactory course in probability theory should be able to use the book for self-study.

It has been my personal experience, as well as my observation of others, that success in dealing with abstract systems rests in large part on the ability to find concrete mental images and constructs which serve as aids in visualizing, remembering, and relating the abstract concepts and ideas. This being so, success in teaching abstract systems depends in similar measure on making explicit use of the most satisfactory images, diagrams and other conceptual aids in the act of communicating ideas.

The literature on probability—both works on pure mathematics and on practical applications—contains a number of such aids to clear thinking, but these aids have not always been exploited fully and efficiently. I can lay little claim to originality in the sense of novelty of ideas or results. Yet I believe the synthesis presented in this book, with its systematic exploitation of several ideas and techniques which ordinarily play only a marginal role in the literature known to me, provides an approach to probability theory which has some definite pedagogical advantages.

Among the features of this presentation which may deserve mention are:

1. A full exploitation of the concept of probability as mass; in particular, the idea that a random variable produces a point-by-point mass transfer from the basic probability space is introduced and utilized in a manner that has proved helpful.

2. Exploitation of minterm maps, minterm expansions, binary designators, and other notions and techniques from the theory of switching, or logic, networks as an aid to systematizing the handling of compound events.

3. Use of the indicator function for events (sets) to provide analytical expressions for discrete-valued random variables.

4. Development of the basic ideas of integration on an abstract space to give unity to the various expressions for mathematical expectation. The mass picture is exploited here in a very significant way to make these ideas comprehensible.

5. Development of a calculus of mathematical expectations which simplifies many arguments that are otherwise burdened with unnecessary details.

None of these ideas or approaches is new. Some have been utilized to good advantage in the literature. What may be claimed is that the systematic exploitation in the manner of this book has provided a treatment of the topic that seems to communicate to my students in a way that no other treatment with which I am familiar has been able to do. This work is written in the hope that this treatment may be equally helpful to others who are seeking a more adequate grasp of this fascinating and powerful subject.

Acknowledgments

My indebtedness to the current literature is so great that it surely must be apparent. I am certainly aware of that debt.

Among the many helpful comments and suggestions by students and colleagues, I am particularly grateful for my discussions with Dr. A. J. Welch and Dr. Shu Lin, graduate students at the time of writing. The critical reviews—from different points of view—by Dr. H. D. Brunk, my former professor, and Dr. John G. Truxal, as well as those by several unnamed reviewers, have aided and stimulated me to produce a better book than I could have written without such counsel.

The patient and dedicated work of Mrs. Arlene McCourt and Mrs. Velma T. Goodwin has been an immeasurable aid in producing a manuscript. I can only hope that the pride and care they showed in their part of the work is matched in some measure by the quality of the contents.

PAUL E. PFEIFFER

Contents

chapter 1

Introduction

As is true of so much of mathematics, probability theory has a long history whose beginnings are largely unknown or obscure. In this chapter we examine very briefly the classical concept of probability which arose in early investigations and which still remains the basis of many applications. It seems reasonably certain that the principal impetus for the development of probability theory came from an interest in games of chance. Interest in gambling is ancient and widespread; games of chance involve an element of “randomness”; it is, in fact, puzzling that the idea of randomness and the attempt to describe it mathematically did not develop earlier. David [1962] discusses this cultural enigma in an interesting study of the early gropings toward a theory of probability and of the early work in the field.

The rudiments of a mathematical theory probably took shape in the sixteenth century. Some evidence of this is provided by a short note written in the early seventeenth century by the famous mathematician and astronomer Galileo Galilei. In a fragment known as *Thoughts about Dice Games*, Galileo dealt with certain problems posed to him by a gambler whose identity is not now known. One of the points of interest in this note (cf. David [1962, pp. 65ff.]) is that Galileo seems to assume that his reader would know how to calculate certain elementary probabilities.

The celebrated correspondence between Blaise Pascal and Pierre de Fermat in 1654, the treatise by Christianus Huygens, 1657, entitled *De Ratiociniis in Aleae Ludo*, and the work *Ars Conjectandi* by James Bernoulli (published posthumously in 1713, but probably written some time about 1690) are landmarks in the formulation and development of the classical theory. The fundamental definition of probability which was accepted in this period, vaguely assumed when not explicitly stated, remained the classical definition until the modern formulations developed in this century provided important extensions and generalizations.

We shall simply examine the classical concept, note some of its limitations, and try to identify the fundamental properties which underlie modern axiomatic formulations. It turns out that the key properties are extremely simple. All the results of the classical theory are obtained as easily in the more general system which we study in this book. Because of the success of the more general system, we shall not examine separately the extensive mathematical system developed upon the classical base. For such a development, one may consult the treatise by Uspensky [1937].

1-1 Basic ideas and the classical definition

The interest in games of chance which stimulated early work in probability not only provided the motivation for that work but also influenced the character of the emerging theory. Almost instinctively, it seems, the best minds attempted to analyze the probability situations into sets of possible outcomes of a gaming operation. These possibilities were then assumed to be "equally likely." The success of the analysis in predicting "chances" led eventually to the precise definition of probability, which remained the classical definition until early in the present century.

Definition 1-1a Classical Probability

A trial is made in which the outcome is one of N equally likely possible outcomes. If, among these N possible outcomes, there are N_A possible outcomes which result in the occurrence of the event A, the *probability of the event A* is defined by

$$P(A) = \frac{N_A}{N}$$

This definition seems to be motivated by two factors:

1. The intuitive idea of *equally likely* possible outcomes
2. The empirical fact of the *statistical regularity of the relative frequencies* of the occurrence of events to be studied

The statistical regularity of the relative frequencies of the occurrence of various events in gambling games has long been observed. In fact, some of the problems posed to noted mathematicians were the result of small variations of observed frequencies from those anticipated by the gamblers (David [1962, pp. 66, 89]). The character of the games was such that the notion of "equally likely" led to successful predictions. So natural did this concept seem that it has been defended vigorously upon philosophical grounds.

Once the definition is made—for whatever reasons—no further appeal to intuition or philosophy is needed. A situation is given; two questions must be answered :

1. How many possible outcomes are there (i.e., what is the value of N)?
2. How many of the possible outcomes result in the occurrence of event A (i.e., what is the value of N_A)?

Once these questions are answered, the probability is determined by the ratio specified in the definition. The problem is to determine answers to these two questions. This, in turn, is a problem of counting the possibilities.

Consider a simple example. Two dice are thrown. Suppose the event A is the event that "a six is thrown." This means that the pair of numbers which appear (in the form of spots) must add to six. What is the probability of throwing a six? First, we must identify the equally likely possible outcomes. Then we must perform the appropriate counting operations. If the dice are "fair," it seems that it is equally likely that any one of the six sides of either of the dice will appear. It is thus natural to consider the appearance of each of the 36 possible pairs of sides of the two dice as equally likely. These various possibilities may be represented simply by pairs of numbers. The first number, being one of the integers 1 through 6, represents the corresponding side of one of the dice. The second number represents the corresponding side of the second die. Thus the number pair (3, 2) indicates the appearance of side 3 of the first die and of side 2 of the second die. The sides are usually numbered according to the number of spots thereon.

We have said there are 36 such pairs. For each possibility on the first die there are six possibilities on the second die. Thus, for the rolling of the dice, we take N to be 36. To determine N_A, we may in this case simply enumerate those pairs for which the sum is 6. These are the pairs (1, 5), (2, 4), (3, 3), (4, 2), and (5, 1). There are five such outcomes, so that $N_A = 5$. The desired probability is thus, *by definition*, $5/36$.

It should not be assumed from the simple example just discussed that probability theory is trivial. Counting, in complex situations, can be a very sophisticated matter, as references to the literature will show (cf. Uspensky [1937] or Feller [1957]). Much of the classical probability theory is devoted to the development of counting techniques. The principal tool is the theory of permutations and combinations. A brief summary of some of the more elementary results is given in Appendix A. An excellent introductory treatment is given in Goldberg [1960, chap. 3]; a more extensive treatment is given in Feller [1957, chaps. 2 through 4].

Upon this simple base a magnificent mathematical structure has been erected. Introduction of the laws of compound probability and of the concepts of conditional probability, random variables, and mathematical expectation have provided a mathematical system rich in content and powerful in its application. As an example of the range of such theory, one should examine a work such as the classical treatise by J. V. Uspensky [1937], entitled *Introduction to Mathematical Probability*. So successful was this development that Uspensky could venture the opinion that modern attempts to provide an axiomatic foundation would result in interesting mental exercises but would have little value for application [*op. cit.*, p. 8].

The classical theory suffers some inherent limitations that inhibit its applications to many problems. Moreover, the success of modern mathematical models in extending the classical theory has provided a more flexible base for applications. Thus it seems desirable, both for applications and for purely mathematical investigations, to move beyond the classical model.

1-2 Motivation for a more general theory

There are two rather obvious limitations of classical probability theory. For one thing, it is limited to situations in which there is only a finite set of possible outcomes. Very simple situations arise, even in classical gambling problems, in which a finite set of possibilities is not adequate. Suppose a game is played until one player is successful in performing a given act (i.e., until he "wins"). Any particular sequence of plays is likely to terminate in a finite number of trials. But there is no a priori assurance that this will happen. A man could conceivably flip a coin indefinitely without ever turning up a head. At any rate, no one can determine a number large enough to include all possible sequences ending in a successful toss. Other simple gaming operations can be conceived in which the game goes on endlessly. In order to account for these possibilities, there must be a model in which the possibilities are not limited to any finite number.

It is also desirable, both for theoretical and practical reasons, to extend the theory to situations in which there is a continuum of possibilities. In such situations, some physical variable may be observed: the height of an individual, the value of an electric current in a wire, the amount of water in a tank, etc. Each of the continuum of possible values of these variables is to be considered a possible outcome.

A second limitation inherent in the classical theory is the assumption of equally likely outcomes. It is noted above that the classical theory seems to be rooted in the two concepts of (1) equally likely outcomes and (2) statistical regularity of relative frequencies. It often occurs that these two concepts do not lead to the same definition. A simple example is the loaded die. For a die which is asymmetrical in terms of mass or shape, it is not intuitively expected that each side will turn up with equal frequency; as a matter of fact, both experience and intuition agree that the relative frequencies will not be the same for the different sides. But *it is expected that the relative frequencies will show statistical regularity*. Experience bears this out in many situations, of which the loaded die is a simple example.

These considerations suggest that the extension of the definition of probability should preserve the essential characteristics of relative frequencies. Two properties prove to be satisfactory for the extension:

1. If f_A is the relative frequency of occurrence of an event A, then $0 \leq f_A \leq 1$.

2. If A and B are mutually exclusive events and C is the event which occurs iffi (if and only if) either A or B occurs, then $f_C = f_A + f_B$.

In the next chapter we begin the development of a theory which defines probability as a function of events; the characteristic properties of the probability function are (1) that it takes values between zero and one and (2) that it has a fundamental *additivity property* for the probability of mutually exclusive events.

The idea of the relative frequency of the occurrence of events plays such an important role in motivating the concept of probability and in interpreting the meaning of the mathematical results that some competent mathematicians have developed mathematical models in which probability is defined as a limit of a relative frequency. This approach has the advantage of tying the fundamental concepts closely to the experiential basis for the introduction of the theoretical model. It has the disadvantage, however, of introducing certain complications into the formulation of the basic definitions and axioms.

It seems far more fruitful to postulate the existence of probabilities which have the simple fundamental properties discussed above. When these probabilities are *interpreted* as relative frequencies, the behavior of the mathematical model can be compared with the behavior of the physical (or other) system that it is intended to represent. The frequency interpretation is aided by the development of certain theorems known under the generic title of *the law of large numbers*. The high degree of correlation between suitable models based on this approach and the observed behavior of many practical systems have provided grounds for confidence in the suitability of such models. This approach is based philosophically on the view that one cannot "prove" anything about the physical world in terms of a mathematical model. One constructs a model, studies its "behavior," uses the results to predict phenomena in the "real world," and evaluates the usefulness of his model in terms of the degree to which the behavior of the mathematical model corresponds to the behavior of the real-world system. "The proof is in the pudding." The growing literature on applications in a wide variety of fields indicates the extent to which such models have been successful (cf. the article by S. S. Shu in Bogdanoff and Kozin [1963] for a brief survey of the history of applications of probability theory in physics and engineering).

Because of these considerations, we do not attempt to examine the theory constructed upon the foundation of the classical definition of probability; instead, we turn immediately to the more general model. Not only does this general theory include the classical theory as a special case; it is often simpler to develop the more general concepts—in spite of certain abstractions—and then examine specific problems from the vantage point provided by this general approach. More elegant solutions and more satisfactory interpretations of problems and solutions are often obtainable with a smaller total effort.

Selected references

DAVID [1962]: "Games, Gods, and Gambling." This interesting work deals with "the origins and history of probability and statistical ideas from the earliest times to the Newtonian era." A readable treatment, with many interesting personal and historical sidelights. The author has a keen interest in the history of ideas as well as in the development of the technical aspects of probability theory in its early stages.

FELLER [1957]: "An Introduction to Probability Theory and Its Applications," vol. 1, 2d ed. An introduction and an extensive treatment of probability theory in the case of a finite or countably infinite number of possible outcomes. Chapters 2, 3, and 4 provide a rather extensive treatment of the problem of counting the ways an event can occur.

GOLDBERG [1960]: "Probability: An Introduction." A lucid treatment of the modern point of view, which is mathematically easy because the author deals only with the case of a finite number of possible outcomes. Chapter 3 provides an excellent introduction to the theory of permutations and combinations needed for many probability problems, both in the classical and in the more general case.

USPENSKY [1937]: "Introduction to Mathematical Probability." A classical treatment of classical probability. This work is still a major reference for many aspects of the mathematical theory and its applications, although its author takes a dim view of the modern axiomatic model which the present work attempts to expound. Available in a paperback edition, it probably should be on the bookshelf of any person having a serious interest in probability theory.

chapter 2

A mathematical model for probability

The discussion in Chap. 1 has shown that the classical theory of probability, based upon a finite set of equally likely possible outcomes of a trial, has severe limitations which make it inadequate for many applications. This is not to dismiss the classical case as trivial, for an extensive mathematical theory and a wide range of applications are based upon this model. It has been possible, by the use of various strategies, to extend the classical case in such a way that the restriction to equally likely outcomes is greatly relaxed. So widespread is the use of the classical model and so ingrained is it in the thinking of those who use it that many people have difficulty in understanding that there can be any other model. In fact, there is a tendency to suppose that one is dealing with physical reality itself, rather than with a model which represents certain aspects of that reality. In spite of this appeal of the classical model, with both its conceptual simplicity and its theoretical power, there are many situations in which it does not provide a suitable theoretical framework for dealing with problems arising in practice. What is needed is a generalization of the notion of probability in a manner that preserves the essential properties of the classical model, but which allows the freedom to encompass a much broader class of phenomena.

In the attempt to develop a more satisfactory theory, we shall seek in a deliberate way to describe a *mathematical model* whose essential features may be correlated with the appropriate features of real-world problems. The history of probability theory (as is true of most theories) is marked both by brilliant intuition and discovery and by confusion and controversy. Until certain patterns had emerged to form the basis of a clear-cut theoretical model, investigators could not formulate problems with precision, and reason about them with mathematical assurance. Long experience was required before the essential patterns were discovered and abstracted. We stand in the fortunate position of having the fruits of this experience distilled in the formulation of a remarkably successful mathematical model.

A mathematical model shares common features with any other type of model. Consider, for example, the type of model, or "mock-up," used extensively in the design of automobiles or aircraft. These models display various essential features: shape, proportion, aerodynamic characteristics, interrelation of certain component parts, etc. Other features, such as weight, details of steering mechanism, and specific materials, may not be incorporated into the particular model used. Such a model is not equivalent to the entity it represents. Its usefulness depends on how well it displays the features it is designed to portray; that is, *its value depends upon how successfully the appropriate features of the model may be related to the "real-life" situation, system, or entity modeled*. To develop a model, one must be aware of its limitations as well as its useful properties.

What we seek, in developing a *mathematical model* of probability, is a *mathematical system* whose concepts and relationships correspond to the appropriate concepts and relationships in the "real world." Once we set up the model (i.e., the mathematical system), we shall study its mathematical behavior in the hope that the patterns revealed in the mathematical system will help in identifying and understanding the corresponding features in real life.

We must be clear about the fact that the mathematical model cannot be used to *prove* anything about the real world, although a study of the model may help us to *discover* important facts about the real world. A model is not true or false; rather, a model fits (i.e., corresponds properly to) or does not fit the real-life situation. A model is useful, or it is not. A model is useful if the three following conditions are met:

1. Problems and situations in the real world can be *translated* into problems and situations in the mathematical model.

2. The model can be studied as a mathematical system to obtain solutions to the *model problems* which are formulated by the translation of real-world problems.

3. The solutions of a model problem can be correlated with or *interpreted* in terms of the corresponding real-world problem.

The mathematical model must be a consistent mathematical system. As such, it has a "life of its own." It may be studied by the mathematician without reference to the translation of real-world problems or the interpretation of its features in terms of real-world counterparts. To be useful from the standpoint of applications, however, not only must it be mathematically sound, but also its results must be physically meaningful when proper interpretation is made. Put negatively, a model is considered unsatisfactory if either (1) the solutions of model problems lead to unrealistic solutions of real-world problems or (2) the model is incomplete or inconsistent mathematically.

Although long experience was needed to produce a satisfactory theory, we need not retrace and relive the mistakes and fumblings which delayed the discovery of an appropriate model. Once the model has been discovered, studied, and refined, it becomes possible for ordinary minds to grasp, in reasonably short time, a pattern which took decades of effort and the insight of genius to develop in the first place.

The most successful model known at present is characterized by considerable mathematical abstractness. A complete study of all the important mathematical questions raised in the process of establishing this system would require a mathematical sophistication and a budget of time and energy not properly to be expected of those whose primary interest is in application (i.e., in solutions to real-world problems). Two facts motivate the study begun in this chapter:

1. Although the details of the mathematics may be sophisticated and difficult, *the central ideas are simple* and *the essential results are often plausible,* even when

difficult to prove.

2. A mastery of the ideas and a reasonable skill in translating real-world problems into model problems make it possible to grasp and solve problems which otherwise are difficult, if not impossible, to solve. *Mastery of this model extends considerably one's ability to deal with real-world problems*.

In addition to developing the fundamental mathematical model, we shall develop certain auxiliary representations which facilitate the grasp of the mathematical model and aid in discovering strategies of solution for problems posed in its terms. We may refer to the combination of these auxiliary representations as the *auxiliary model*.

Although the primary goal of this study is the ability to solve real-world problems, success in achieving this goal requires a reasonable mastery of the mathematical model and of the strategies and techniques of solution of problems posed in terms of this model. Thus considerable attention must be given to the model itself. As we have already noted, the model may be studied as a thing in itself, with a "life of its own." This means that we shall be engaged in developing a mathematical theory. The study of this mathematics can be an interesting and challenging game in itself, with important dividends in training in analytical thought. At times we must be content to play the game, until a stage is reached at which we may attempt a new correlation of the model with the real world. But as we reach these points in the development of the theory, repeated success in the act of interpretation will serve to increase our confidence in the model and to make it easier to comprehend its character and see its implications for the real world.

The model to be developed is essentially the axiomatic system described by the mathematician A. N. Kolmogorov (1903–   ), who brought together in a classical monograph [1933] many streams of development. This monograph is now available in English translation under the title *Foundations of the Theory of Probability* [1956]. The Kolmogorov model presents mathematical probability as a special case of abstract measure theory. Our exposition utilizes some concrete but essentially sound representations to aid in grasping the abstract concepts and relations of this model. We present the concepts and their relations with considerable precision, although we do not always attempt to give the most general formulation. At many places we borrow mathematical theorems without proof. We sometimes note critical questions without making a detailed examination; we merely indicate how they have been resolved. Emphasis is on concepts, content of theorems, interpretations, and strategies of problem solution suggested by a grasp of the essential content of the theory. Applications emphasize the translation of physical assumptions into statements involving the precise concepts of the mathematical model.

It is assumed in this chapter that the reader is reasonably familiar with the elements of set theory and the elementary operations with sets. Adequate treatments of this material are readily available in the literature. A sketch of some of these ideas is given in Appendix B, for ready reference. Some specialized results, which have been developed largely in connection with the application of set theory and boolean algebra to switching circuits, are summarized in Sec. 2–6. A number of references for supplementary reading are listed at the end of this chapter.

Sets, Events, and Switching [1964]. A number of references for supplementary reading are listed at the end of this chapter.

2-1 In search of a model

The discussion in the previous introductory paragraphs has indicated that, to establish a mathematical model, we must first identify the significant concepts, patterns, relations, and entities in the "real world" which we wish to represent. Once these features are identified, we must seek appropriate mathematical counterparts. These mathematical counterparts involve concepts and relations which must be defined or postulated and given appropriate names and symbolic representations.

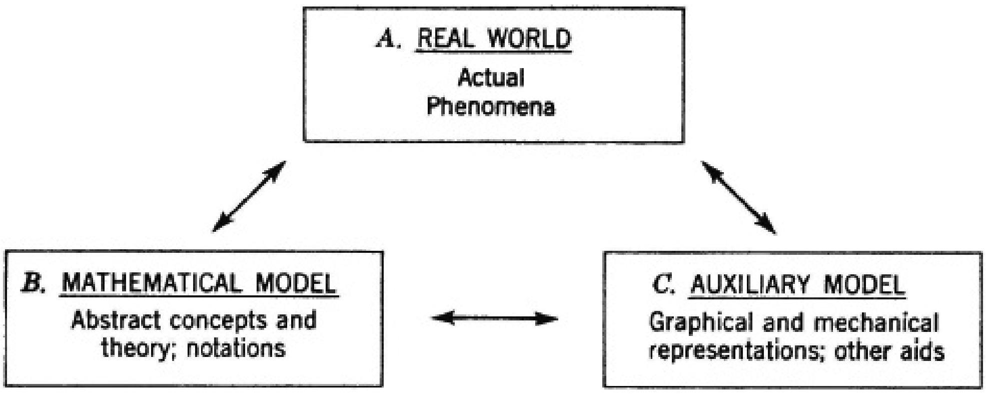

Fig. 2-1-1 Diagrammatic representation of the relationships between the "real world" and the models.

In order to be clear about the situation that exists when we utilize mathematical models, let us make a diagrammatic representation as in Fig. 2-1-1. In this diagram, we analyze the object of our investigation into three component parts:

A. The *real world* of actual phenomena, known to us through the various means of experiencing these phenomena.

B. The imaginary world of the *mathematical model*, with its abstract concepts and theory. An important feature of this model is the use of symbolic notational schemes which enable us to state relationships and facts with great precision and economy.

C. An *auxiliary model*, consisting of various graphical, mechanical, and other aids to visualization, remembering, and even discovering important features about the mathematical model. It seems likely that even the purest of mathematicians, dealing with the most abstract mathematical systems, employ, consciously or unconsciously, concrete mental images as the carriers of their thought patterns. We shall develop explicitly some concrete representations to aid in thinking about the abstract mathematical model; these in turn will help us to think clearly and systematically about the patterns abstracted from (i.e., lifted out of) our experience of the real world of phenomena.

Much of our attention and effort will be devoted to establishing the mathematical model *B* and to a study of its characteristics. In doing this, we shall be concerned to relate the various aspects of the mathematical model to corresponding aspects of the auxiliary model *C*, as an aid to learning and remembering the important characteristics of the mathematical model. Our real goal as engineers and scientists, however, is to use our knowledge of the mathematical model as an aid in dealing with problems in the real world. This means that we must be able to move from one part of our system to another with freedom and insight. For clarity and emphasis, we may find it helpful to indicate the important transitions in the following manner:

A → *B: Translation* of real-world concepts, relations, and problems into terms of the concepts of the mathematical model.

B → *A: Interpretation* of the mathematical concepts and results in terms of real-world phenomena. This may be referred to as the *primary interpretation*.

B → *C: Interpretation* of the mathematical concepts and results in terms of various concrete representations (mass picture, mapping concepts, etc.). This may be referred to as a *secondary interpretation*.

C → *B*: The movement from the auxiliary model to the mathematical model exploits the concrete imagery of the former to aid in discovering new results, remembering and extending previously discovered results, and evolving strategies for the solution of model problems.

A ↔ *C*: The correlation of features in *A* and *C* often aids both the translation of real-world problems into model problems and the interpretation of the mathematical results. In other words, the best path from *A* to *B* or from *B* to *A* may be through *C*.

The first element to be modeled is the relative frequency of the occurrence of an event. It is an *empirical fact* that in many investigations the relative frequencies of occurrence of various events exhibit a remarkable *statistical regularity* when a large number of trials are made. This feature of many games of chance served (as we noted in Chap. 1) to motivate much of the early development of probability theory. In fact, many of the questions posed by gamblers to the mathematicians of their time were evoked by the fact that observed frequencies deviated slightly from that which they expected.

This phenomenon of constant relative frequencies is in no way limited to games of chance. Modern statistical communication theory, for example, makes considerable use of the remarkable statistical regularities which characterize languages. The relative frequencies of occurrence of symbols of the alphabet (including all symbols such as space, numbers, punctuation, etc.), of symbol pairs, triples, etc., and of words, word combinations, etc., have been studied extensively and are known to be quite stable. Of course, exceptions are known. For example, Pierce [1961] quotes a paragraph from a novel which is written without the use of a single letter *E* in its entire 267 pages. In ordinary English, the letter *E* is used more frequently than any other letter of the alphabet. Such marked deviations from the usual patterns require special effort or indicate unusual situations. In the normal course of affairs one may expect rather close adherence to the common patterns of statistical regularity.

The whole life insurance industry is based upon so-called mortality tables, which predict the relative frequency of deaths at various ages (or give information equivalent to specifying these frequencies). In a similar manner, reliability engineering makes extensive use of the life expectancy of articles manufactured by a given process. These life expectancies are based on the equivalent of mortality tables for the products manufactured. Closely related is the theory of errors, in which the relative frequencies of errors of various magnitudes may be assumed to follow quite closely certain well-known patterns.

The empirical fact of a stable (i.e., constant) relative frequency serves as the basis in experience for choosing a mathematical model of probability. This empirical fact cannot in any way imply logically the existence of a probability number. We set up a model by *postulating* the existence of an ideal number, which we call the *probability* of an event. If this is a sound model and we have chosen the number properly, we may expect the relative frequency of occurrence of the event in a large number of trials to lie quite close to this number. But we cannot "prove" the existence in the real world of such an ideal limiting frequency. We simply set up a model. We then examine the "behavior" of the model, *interpret* the probabilities as relative frequencies, check or *test* these probabilities against observed frequencies (where possible), and *try to determine experimentally whether the model fits the real-world situation to be modeled. We can build up confidence in our model; we cannot prove or disprove our model*. If it works for enough different problems, we move with considerable confidence to new problems. Continued success tends to increase our confidence, so that we come to place a large degree of reliance upon the adequacy of the model to predict behavior in the real world.

If the feature of the real world to be represented by probability is the relative frequency of occurrence of a given event in a large number of trials, then the probability must be a number associated with this event. These numbers must obey the same laws as relative frequencies. We can readily list several elementary properties which must therefore be possessed by the probability model.

1. If an event is sure to occur, its relative frequency, and hence its probability, must be unity. Similarly, if an event cannot possibly occur, its probability must be 0.

2. Probabilities are real numbers, lying between 0 and 1.

3. If two events are mutually exclusive (i.e., cannot both happen on any one trial), the probability of the compound event that either one or the other of the original events will occur is the sum of the individual probabilities.

4. The probability that an event will not occur is 1 minus the probability of the occurrence of the event.

5. If the occurrence of one event implies the occurrence of a second event, the relative frequency of occurrence of the second event must be at least as great as that of the first event. Thus the probability of the second event must be at least as great as that of the first event.

Many more such properties could be enumerated. One of the concerns of the model maker is to discover the most basic list, in the sense that the properties included in this list imply as logical consequences all the other desirable or necessary properties. When we come to the formal presentation of our mathematical model in Sec. 2-3, we shall see that the basic list desired is contained in the list of properties above.

In order to realize an economy of expression, we shall need to introduce an