Applied Probability

By:
Paul E Pfeiffer

Applied Probability

By:
Paul E Pfeiffer

Online:
< http://cnx.org/content/col10708/1.6/ >

C O N N E X I O N S

Rice University, Houston, Texas

Collection structure revised: August 31, 2009
PDF generated: October 26, 2012
For copyright and attribution information for the modules contained in this collection, see p. 618.

Table of Contents

Preface to Pfeiffer Applied Probability[1]

The course

This is a "first course" in the sense that it presumes no previous course in probability. The units are modules taken from the unpublished text: Paul E. Pfeiffer, ELEMENTS OF APPLIED PROBABILITY, USING MATLAB. The units are numbered as they appear in the text, although of course they may be used in any desired order. For those who wish to use the order of the text, an outline is provided, with indication of which modules contain the material.

The mathematical prerequisites are ordinary calculus and the elements of matrix algebra. A few standard series and integrals are used, and double integrals are evaluated as iterated integrals. The reader who can evaluate simple integrals can learn quickly from the examples how to deal with the iterated integrals used in the theory of expectation and conditional expectation. Appendix B (Section 17.2) provides a convenient compendium of mathematical facts used frequently in this work. And the symbolic toolbox, implementing MAPLE, may be used to evaluate integrals, if desired.

In addition to an introduction to the essential features of basic probability in terms of a precise mathematical model, the work describes and employs user defined MATLAB procedures and functions (which we refer to as *m-programs*, or simply *programs*) to solve many important problems in basic probability. This should make the work useful as a stand alone exposition as well as a supplement to any of several current textbooks.

Most of the programs developed here were written in earlier versions of MATLAB, but have been revised slightly to make them quite compatible with MATLAB 7. In a few cases, alternate implementations are available in the Statistics Toolbox, but are implemented here directly from the basic MATLAB program, so that students need only that program (and the symbolic mathematics toolbox, if they desire its aid in evaluating integrals).

Since machine methods require precise formulation of problems in appropriate mathematical form, it is necessary to provide some supplementary analytical material, principally the so-called *minterm analysis*. This material is not only important for computational purposes, but is also useful in displaying some of the structure of the relationships among events.

A probability model

Much of "real world" probabilistic thinking is an amalgam of intuitive, plausible reasoning and of statistical knowledge and insight. Mathematical probability attempts to to lend precision to such probability analysis by employing a suitable *mathematical model*, which embodies the central underlying principles and structure. A successful model serves as an aid (and sometimes corrective) to this type of thinking.

Certain concepts and patterns have emerged from experience and intuition. The mathematical formulation (the mathematical model) which has most successfully captured these essential ideas is rooted in measure theory, and is known as the *Kolmogorov model*, after the brilliant Russian mathematician A.N. Kolmogorov (1903-1987).

[1]This content is available online at <http://cnx.org/content/m23242/1.8/>.

One cannot prove that a model is *correct*. Only experience may show whether it is *useful* (and not incorrect). The usefulness of the Kolmogorov model is established by examining its structure and showing that patterns of uncertainty and likelihood in any practical situation can be represented adequately. Developments, such as in this course, have given ample evidence of such usefulness.

The most fruitful approach is characterized by an interplay of

- A formulation of the problem in precise terms of the model and careful mathematical analysis of the problem so formulated.
- A grasp of the problem based on experience and insight. This underlies both problem formulation and interpretation of analytical results of the model. Often such insight suggests approaches to the analytical solution process.

MATLAB: A tool for learning

In this work, we make extensive use of MATLAB as an aid to analysis. I have tried to write the MATLAB programs in such a way that they constitute useful, ready-made tools for problem solving. Once the user understands the problems they are designed to solve, the solution strategies used, and the manner in which these strategies are implemented, the collection of programs should provide a useful resource.

However, my primary aim in exposition and illustration is to *aid the learning process* and to deepen insight into the structure of the problems considered and the strategies employed in their solution. Several features contribute to that end.

1. Application of machine methods of solution requires precise formulation. The data available and the fundamental assumptions must be organized in an appropriate fashion. The requisite *discipline* for such formulation often contributes to enhanced understanding of the problem.
2. The development of a MATLAB program for solution requires careful attention to possible solution strategies. One cannot instruct the machine without a clear grasp of what is to be done.
3. I give attention to the tasks performed by a program, with a general description of how MATLAB carries out the tasks. The reader is not required to trace out all the programming details. However, it is often the case that available MATLAB resources suggest alternative solution strategies. Hence, for those so inclined, attention to the details may be fruitful. I have included, as a separate collection, the m-files written for this work. These may be used as patterns for extensions as well as programs in MATLAB for computations. Appendix A (Section 17.1) provides a directory of these m-files.
4. Some of the details in the MATLAB script are presentation details. These are refinements which are not essential to the solution of the problem. But they make the programs more readily usable. And they provide illustrations of MATLAB techniques for those who may wish to write their own programs. I hope many will be inclined to go beyond this work, modifying current programs or writing new ones.

An Invitation to Experiment and Explore

Because the programs provide considerable freedom from the burden of computation and the tyranny of tables (with their limited ranges and parameter values), standard problems may be approached with a new spirit of experiment and discovery. When a program is selected (or written), it embodies one method of solution. There may be others which are readily implemented. The reader is invited, even urged, to explore! The user may experiment to whatever degree he or she finds useful and interesting. The possibilities are endless.

Acknowledgments

After many years of teaching probability, I have long since lost track of all those authors and books which have contributed to the treatment of probability in this work. I am aware of those contributions and am most eager to acknowledge my indebtedness, although necessarily without specific attribution.

The power and utility of MATLAB must be attributed to to the long-time commitment of Cleve Moler, who made the package available in the public domain for several years. The appearance of the professional versions, with extended power and improved documentation, led to further appreciation and utilization of its potential in applied probability.

The Mathworks continues to develop MATLAB and many powerful "tool boxes," and to provide leadership in many phases of modern computation. They have generously made available MATLAB 7 to aid in checking for compatibility the programs written with earlier versions. I have not utilized the full potential of this version for developing professional quality user interfaces, since I believe the simpler implementations used herein bring the student closer to the formulation and solution of the problems studied.

CONNEXIONS

The development and organization of the CONNEXIONS modules has been achieved principally by two people: C.S.(Sid) Burrus a former student and later a faculty colleague, then Dean of Engineering, and most importantly a long time friend; and Daniel Williamson, a music major whose keyboard skills have enabled him to set up the text (especially the mathematical expressions) with great accuracy, and whose dedication to the task has led to improvements in presentation. I thank them and others of the CONNEXIONS team who have contributed to the publication of this work.

Paul E. Pfeiffer

Rice University

Chapter 1

Probability Systems

1.1 Likelihood[1]

1.1.1 Introduction

Probability models and techniques permeate many important areas of modern life. A variety of types of random processes, reliability models and techniques, and statistical considerations in experimental work play a significant role in engineering and the physical sciences. The solutions of management decision problems use as aids decision analysis, waiting line theory, inventory theory, time series, cost analysis under uncertainty — all rooted in applied probability theory. Methods of statistical analysis employ probability analysis as an underlying discipline.

Modern probability developments are increasingly sophisticated mathematically. To utilize these, the practitioner needs a sound conceptual basis which, fortunately, can be attained at a moderate level of mathematical sophistication. There is need to develop a feel for the structure of the underlying mathematical model, for the role of various types of assumptions, and for the principal strategies of problem formulation and solution.

Probability has roots that extend far back into antiquity. The notion of "chance" played a central role in the ubiquitous practice of gambling. But chance acts were often related to magic or religion. For example, there are numerous instances in the Hebrew Bible in which decisions were made "by lot" or some other chance mechanism, with the understanding that the outcome was determined by the will of God. In the New Testament, the book of Acts describes the selection of a successor to Judas Iscariot as one of "the Twelve." Two names, Joseph Barsabbas and Matthias, were put forward. The group prayed, then drew lots, which fell on Matthias.

Early developments of probability as a mathematical discipline, freeing it from its religious and magical overtones, came as a response to questions about games of chance played repeatedly. The mathematical formulation owes much to the work of Pierre de Fermat and Blaise Pascal in the seventeenth century. The game is described in terms of a well defined trial (a play); the result of any trial is one of a specific set of distinguishable outcomes. Although the result of any play is not predictable, certain "statistical regularities" of results are observed. The possible results are described in ways that make each result seem equally likely. If there are N such possible "equally likely" results, each is assigned a probability $1/N$.

The developers of mathematical probability also took cues from early work on the analysis of statistical data. The pioneering work of John Graunt in the seventeenth century was directed to the study of "vital statistics," such as records of births, deaths, and various diseases. Graunt determined the fractions of people in London who died from various diseases during a period in the early seventeenth century. Some thirty years later, in 1693, Edmond Halley (for whom the comet is named) published the first life insurance tables. To apply these results, one considers the selection of a member of the population on a chance basis. One

[1] This content is available online at <http://cnx.org/content/m23243/1.8/>.

then assigns the probability that such a person will have a given disease. The trial here is the selection of a person, but the interest is in certain characteristics. We may speak of the event that the person selected will die of a certain disease– say "consumption." Although it is a person who is selected, it is death from consumption which is of interest. Out of this statistical formulation came an interest not only in probabilities as fractions or relative frequencies but also in averages or expectatons. These averages play an essential role in modern probability.

We do not attempt to trace this history, which was long and halting, though marked by flashes of brilliance. Certain concepts and patterns which emerged from experience and intuition called for clarification. We move rather directly to the mathematical formulation (the "mathematical model") which has most successfully captured these essential ideas. This is the model, rooted in the mathematical system known as measure theory, is called the *Kolmogorov model*, after the brilliant Russian mathematician A.N. Kolmogorov (1903-1987). Kolmogorov succeeded in bringing together various developments begun at the turn of the century, principally in the work of E. Borel and H. Lebesgue on measure theory. Kolmogorov published his epochal work in German in 1933. It was translated into English and published in 1956 by Chelsea Publishing Company.

1.1.2 Outcomes and events

Probability applies to situations in which there is a well defined trial whose possible outcomes are found among those in a given basic set. The following are typical.

- A pair of dice is rolled; the outcome is viewed in terms of the numbers of spots appearing on the top faces of the two dice. If the outcome is viewed as an ordered pair, there are thirty six equally likely outcomes. If the outcome is characterized by the total number of spots on the two die, then there are eleven possible outcomes (not equally likely).
- A poll of a voting population is taken. Outcomes are characterized by responses to a question. For example, the responses may be categorized as positive (or favorable), negative (or unfavorable), or uncertain (or no opinion).
- A measurement is made. The outcome is described by a number representing the magnitude of the quantity in appropriate units. In some cases, the possible values fall among a finite set of integers. In other cases, the possible values may be any real number (usually in some specified interval).
- Much more sophisticated notions of outcomes are encountered in modern theory. For example, in communication or control theory, a communication system experiences only one signal stream in its life. But a communication system is not designed for a single signal stream. It is designed for one of an infinite *set* of possible signals. The likelihood of encountering a certain kind of signal is important in the design. Such signals constitute a subset of the larger set of all possible signals.

These considerations show that our probability model must deal with

- A *trial* which results in (selects) an *outcome* from a *set* of conceptually possible outcomes. The trial is not successfully completed until one of the outcomes is realized.
- Associated with each outcome is a certain characteristic (or combination of characteristics) pertinent to the problem at hand. In polling for political opinions, it is a person who is selected. That person has many features and characteristics (race, age, gender, occupation, religious preference, preferences for food, etc.). But the primary feature, which characterizes the outcome, is the political opinion on the question asked. Of course, some of the other features may be of interest for analysis of the poll.

Inherent in informal thought, as well as in precise analysis, is the notion of an *event* to which a *probability* may be assigned as a measure of the *likelihood* the event will *occur* on any trial. A successful mathematical model must formulate these notions with precision. An *event* is identified in terms of the characteristic of the outcome observed. The event "a favorable response" to a polling question *occurs* if the outcome observed has that characteristic; i.e., iff (if and only if) the respondent replies in the affirmative. A hand of five cards is drawn. The event "one or more aces" *occurs* iff the hand actually drawn has at least one ace. If that same hand has two cards of the suit of clubs, then the event “two clubs” has *occurred*. These considerations lead to the following definition.

Definition. The *event* determined by some characteristic of the possible outcomes is the set of those outcomes having this characteristic. The event *occurs* iff the outcome of the trial is a member of that set (i.e., has the characteristic determining the event).

- The event of throwing a “seven” with a pair of dice (which we call the event SEVEN) consists of the set of those possible outcomes with a total of seven spots turned up. The event SEVEN occurs iff the outcome is one of those combinations with a total of seven spots (i.e., belongs to the event SEVEN). This could be represented as follows. Suppose the two dice are distinguished (say by color) and a picture is taken of each of the thirty six possible combinations. On the back of each picture, write the number of spots. Now the event SEVEN consists of the set of all those pictures with seven on the back. Throwing the dice is equivalent to selecting randomly one of the thirty six pictures. The event SEVEN occurs iff the picture selected is one of the set of those pictures with seven on the back.
- Observing for a very long (theoretically infinite) time the signal passing through a communication channel is equivalent to selecting one of the conceptually possible signals. Now such signals have many characteristics: the maximum peak value, the frequency spectrum, the degree of differentibility, the average value over a given time period, etc. If the signal has a peak absolute value less than ten volts, a frequency spectrum essentially limited from 60 herz to 10,000 herz, with peak rate of change 10,000 volts per second, then it is *one* of the *set* of signals with those characteristics. The event "the signal has these characteristics" has occured. This set (event) consists of an uncountable infinity of such signals.

One of the advantages of this formulation of an event as a subset of the basic set of possible outcomes is that we can use elementary set theory as an aid to formulation. And tools, such as Venn diagrams and indicator functions (Section 1.3) for studying event combinations, provide powerful aids to establishing and visualizing relationships between events. We formalize these ideas as follows:

- Let Ω be the set of all possible outcomes of the basic trial or experiment. We call this the *basic space* or the *sure event*, since if the trial is carried out successfully the outcome will be in Ω; hence, the event Ω is sure to occur on any trial. We must specify unambiguously what outcomes are “possible.” In flipping a coin, the only accepted outcomes are “heads” and “tails.” Should the coin stand on its edge, say by leaning against a wall, we would ordinarily consider that to be the result of an improper trial.
- As we note above, each outcome may have several characteristics which are the basis for describing events. Suppose we are drawing a single card from an ordinary deck of playing cards. Each card is characterized by a “face value” (two through ten, jack, queen, king, ace) and a “suit” (clubs, hearts, diamonds, spades). An ace is drawn (the event ACE occurs) iff the outcome (card) belongs to the set (event) of four cards with ace as face value. A heart is drawn iff the card belongs to the set of thirteen cards with heart as suit. Now it may be desirable to specify events which involve various logical combinations of the characteristics. Thus, we may be interested in the event the face value is jack *or* king *and* the suit is heart *or* spade. The set for jack *or* king is represented by the union $J \cup K$ and the set for heart *or* spade is the union $H \cup S$. The occurrence of both conditions means the outcome is in the intersection (common part) designated by $\cap$. Thus the event referred to is

$$E = (J \cup K) \cap (H \cup S) \tag{1.1}$$

  The notation of set theory thus makes possible a precise formulation of the event E.
- Sometimes we are interested in the situation in which the outcome does *not* have one of the characteristics. Thus the set of cards which does not have suit heart is the set of all those outcomes not in event H. In set theory, this is the *complementary* set (event) H^c.
- Events are *mutually exclusive* iff not more than one can occur on any trial. This is the condition that the sets representing the events are disjoint (i.e., have no members in common).
- The notion of the *impossible event* is useful. The impossible event is, in set terminology, the *empty set*$\varnothing$. Event $\varnothing$ cannot occur, since it has no members (contains no outcomes). One use of $\varnothing$ is to

provide a simple way of indicating that two sets are mutually exclusive. To say $AB = \varnothing$ (here we use the alternate AB for $A \cap B$) is to assert that events A and B have no outcome in common, hence cannot both occur on any given trial.
- Set *inclusion* provides a convenient way to designate the fact that event A implies event B, in the sense that the occurrence of A requires the occurrence of B. The set relation $A \subset B$ signifies that every element (outcome) in A is also in B. If a trial results in an outcome in A (event A occurs), then that outcome is also in B (so that event B has occurred).

The language and notaton of sets provide a precise language and notation for events and their combinations. We collect below some useful facts about logical (often called Boolean) combinations of events (as sets). The notion of Boolean combinations may be applied to arbitrary classes of sets. For this reason, it is sometimes useful to use an *index set* to designate membership. We say the index J is *countable* if it is finite or countably infinite; otherwise it is *uncountable*. In the following it may be arbitrary.

$$\{A_i : i \in J\} \quad \text{is the class of sets } A_i\text{, one for each index } i \text{ in the index set } J \tag{1.2}$$

For example, if $J = \{1, 2, 3\}$ then $\{A_i : i \in J\}$ is the class $\{A_1, A_2, A_3\}$, and

$$\bigcup_{i \in J} A_i = A_1 \cup A_2 \cup A_3, \qquad \bigcap_{i \in J} A_i = A_1 \cap A_2 \cap A_3, \tag{1.3}$$

If $J = \{1, 2, \cdots\}$ then $\{A_i : i \in J\}$ is the sequence $\{A_1 : 1 \le i\}$. and

$$\bigcup_{i \in J} A_i = \bigcup_{i=1}^{\infty} A_i, \qquad \bigcap_{i \in J} A_i = \bigcap_{i=1}^{\infty} A_i \tag{1.4}$$

If event E is the *union* of a class of events, then event E occurs iff *at least one* event in the class occurs. If F is the *intersection* of a class of events, then event F occurs iff *all* events in the class occur on the trial.

The role of disjoint unions is so important in probability that it is useful to have a symbol indicating the union of a disjoint class. We use the big V to indicate that the sets combined in the union are disjoint. Thus, for example, we write

$$A = \bigvee_{i=1}^{n} A_i \quad \text{to signify} \quad A = \bigcup_{i=1}^{n} A_i \quad \text{with the proviso that the } A_i \text{ form a disjoint class} \tag{1.5}$$

Example 1.1: Events derived from a class

Consider the class $\{E_1, E_2, E_3\}$ of events. Let A_k be the event that exactly k occur on a trial and B_k be the event that k or more occur on a trial. Then

$$A_0 = E_1^c E_2^c E_3^c, \quad A_1 = E_1 E_2^c E_3^c \bigvee E_1^c E_2 E_3^c \bigvee E_1^c E_2^c E_3 \quad A_2 = E_1 E_2 E_3^c \bigvee E_1 E_2^c E_3 \bigvee E_1^c E_2 E_3, \quad A_3 = E_1 E_2 E_3 \tag{1.6}$$

The unions are disjoint since each pair of terms has E_i in one and ${E_i}^c$ in the other, for at least one i. Now the B_k can be expressed in terms of the A_k. For example

$$B_2 = A_2 \bigvee A_3 \tag{1.7}$$

The union in this expression for B_2 is disjoint since we cannot have exactly two of the E_i occur *and* exactly three of them occur on the same trial. We may express B_2 directly in terms of the E_i as follows:

$$B_2 = E_1 E_2 \cup E_1 E_3 \cup E_2 E_3 \tag{1.8}$$

Here the union is *not* disjoint, in general. However, if one pair, say $\{E_1, E_3\}$ is disjoint, then $E_1E_3 = \emptyset$ and the pair $\{E_1E_2, E_2E_3\}$ is disjoint (draw a Venn diagram). Suppose C is the event the first two occur or the last two occur but no other combination. Then

$$C = E_1E_2E_3^c \bigvee E_1^cE_2E_3 \tag{1.9}$$

Let D be the event that one or three of the events occur.

$$D = A_1 \bigvee A_3 = E_1E_2^cE_3^c \bigvee E_1^cE_2E_3^c \bigvee E_1^cE_2^cE_3 \bigvee E_1E_2E_3 \tag{1.10}$$

Two important patterns in set theory known as *DeMorgan's rules* are useful in the handling of events. For an arbitrary class $\{A_i : i \in J\}$ of events,

$$\left[\bigcup_{i \in J} A_i\right]^c = \bigcap_{i \in J} A_i^c \quad \text{and} \quad \left[\bigcap_{i \in J} A_i\right]^c = \bigcup_{i \in J} A_i^c \tag{1.11}$$

An outcome is not in the union (i.e., not in at least one) of the A_i iff it fails to be in all A_i, and it is not in the intersection (i.e. not in all) iff it fails to be in at least one of the A_i.

Example 1.2: Continuation of Example 1.1 (Events derived from a class)
Express the event of no more than one occurrence of the events in $\{E_1, E_2, E_3\}$ as $B_2{}^c$.

$$B_2^c = [E_1E_2 \cup E_1E_3 \cup E_2E_3]^c = (E_1^c \cup E_2^c)(E_1^c \cup E_3^c)\left(E_2^3E_3^c\right) = E_1^cE_2^c \cup E_1^cE_3^c \cup E_2^cE_3^c \tag{1.12}$$

The last expression shows that not more than one of the E_i occurs iff at least two of them fail to occur.

1.2 Probability Systems[2]

1.2.1 Probability measures

In the module "Likelihood" (Section 1.1) we introduce the notion of a basic space Ω of all possible outcomes of a trial or experiment, events as subsets of the basic space determined by appropriate characteristics of the outcomes, and logical or Boolean combinations of the events (unions, intersections, and complements) corresponding to logical combinations of the defining characteristics.

Occurrence or nonoccurrence of an event is determined by characteristics or attributes of the outcome observed on a trial. Performing the trial is visualized as selecting an outcome from the basic set. An event occurs whenever the selected outcome is a member of the subset representing the event. As described so far, the selection process could be quite deliberate, with a prescribed outcome, or it could involve the uncertainties associated with "chance." Probability enters the picture only in the latter situation. Before the trial is performed, there is *uncertainty* about which of these latent possibilities will be realized. *Probability* traditionally is a number assigned to an event indicating the *likelihood* of the occurrence of that event on any trial.

We begin by looking at the *classical* model which first successfully formulated probability ideas in mathematical form. We use modern terminology and notation to describe it.

Classical probability

1. The basic space Ω consists of a finite number N of possible outcomes.
    - There are thirty six possible outcomes of throwing two dice.

[2]This content is available online at <http://cnx.org/content/m23244/1.8/>.