

1.1: Likelihood

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 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.

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 differentiability, 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 occurred. 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](#) 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

$$E = (J \cup K) \cap (H \cup S)$$

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* \emptyset . Event \emptyset cannot occur, since it has no members (contains no outcomes). One use of \emptyset is to provide a simple way of indicating that two sets are mutually exclusive. To say $AB = \emptyset$ (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.
- The language and notation 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.

$A_i : i \in J$ is the class of sets A_i , one for each index i in the index set J

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

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

If $J = 1, 2, \dots$ then $A_i : i \in J$ is the sequence $A_1 : 1 \leq i$, and

$$\bigcup_{i \in J} A_i = \bigcup_{i=1}^{\infty} A_i, \quad \bigcap_{i \in J} A_i = \bigcap_{i=1}^{\infty} A_i$$

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 \bigvee to indicate that the sets combined in the union are disjoint. Thus, for example, we write

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

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

$$A_0 = E_1^c E_2^c E_3^c, A_1 = E_1 E_2^c E_3^c \vee E_1^c E_2 E_3^c \vee E_1^c E_2^c E_3, A_2 = E_1 E_2 E_3^c \vee E_1 E_2^c E_3 \vee E_1^c E_2 E_3, A_3 = E_1 E_2 E_3$$

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

$$B_2 = A_2 \vee A_3$$

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:

$$B_2 = E_1 E_2 \cup E_1 E_3 \cup E_2 E_3$$

Here the union is not disjoint, in general. However, if one pair, say E_1, E_3 is disjoint, then $E_1 E_3 = \emptyset$ and the pair $E_1 E_2, E_2 E_3$ 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

$$C = E_1 E_2 E_3^c \vee E_1^c E_2 E_3$$

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

$$D = A_1 \vee A_3 = E_1 E_2^c E_3^c \vee E_1^c E_2 E_3^c \vee E_1^c E_2^c E_3 \vee E_1 E_2 E_3$$

The 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,

$$[\bigcup_{i \in J} A_i]^c = \bigcap_{i \in J} A_i^c \text{ and } [\bigcap_{i \in J} A_i]^c = \bigcup_{i \in J} A_i^c$$

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 .

continuation of example

Express the event of no more than one occurrence of the events in E_1, E_2, E_3 as B_2^c .

$$B_2^c = [E_1 E_2 \cup E_1 E_3 \cup E_2 E_3]^c = (E_1^c \cup E_2^c)(E_1^c \cup E_3^c)(E_2^c \cup E_3^c) = E_1^c E_2^c \cup E_1^c E_3^c \cup E_2^c E_3^c$$

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

This page titled [1.1: Likelihood](#) is shared under a [CC BY 3.0](#) license and was authored, remixed, and/or curated by [Paul Pfeiffer](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.