FUNDAMENTALS OF PROBABILITY THEORY

M. E. Harr, Purdue University

The components of a pavement system, its loadings and responses, its constitutive materials, and conditions of weather vary in time and location in a random manner. Mathematical models of such systems are known as stochastic processes. This paper presents some fundamentals of probability theory that form the building blocks of such processes. Specific topics treated are deterministic and stochastic systems, randomness and probability, tree diagrams, permutations and computations, conditional probabilities, independence, and Bayes' theorem. Examples are presented to demonstrate the use of the concepts relative to factors entering the analysis, design, construction, and proofing of pavement systems. The concept of chance as it applies to dice or cards is discussed. In this paper a collection of tools is described, and their use is demonstrated.

•ACHIEVEMENTS in transportation technology during the last 2 decades have increased the need for pavement evaluation procedures with which to assess the future trends of pavement behavior. The rates and magnitudes of loadings imposed on today's pavement systems surpass those previously experienced, especially those due to air transport vehicles. The nature of these loadings places greater demands on pavements than they were designed and constructed for. The deterioration of today's pavements has become a major problem of civil engineering.

The problem facing the profession today is not how to design and build new pavement systems for greater frequency and magnitude of loadings, but how to upgrade and provide the remedial measures for existing pavement systems to meet current and future traffic demands.

In pavement design and analysis, factors that are commonly swept under the carpet in other analytical problem areas cannot be assumed away. For example, a pavement consists of distinct layers with unknown contacts at their interface; the layers may or may not be in contact in space or in time. Imposed loadings (wheels) are relatively large in area compared with the thickness of the surface layer; consequently, Saint-Venant's principle cannot be invoked to change the system to an equivalent homogeneous and isotropic body. Ambient conditions greatly alter the properties of the layers, which range from thermal plastic, temperature-sensitive materials to granular soils whose actions depend greatly on their voids. Each layer is composed of complex conglomerations of discrete particles of varying shapes, sizes, and orientations. In addition, loads are variable in both magnitude and time and are dynamic in nature. It is not surprising how poor predictions of the transmission of induced energy through such systems have been. Randomness alone dictates the probabilistic (casual) rather than deterministic (causal) treatment.

DETERMINISTIC AND STOCHASTIC SYSTEMS

Systems that can be described by unique explicit mathematical relationships are said to be deterministic. An example of a common deterministic system is shown in Figure 1a. The system is composed of the mass m, suspended from a linear spring (with spring constant k), which is displaced an amount Δ from its equilibrium position. The mathematical relationship of the response y(t) is

$$y(t) = \Delta \cos \sqrt{\frac{k}{m}}\, t \tag{1}$$

for $t \geq 0$. This expression provides the unique position of the mass y(t) at any instant of time; hence, the system is completely determined or deterministic. A pictorial representation of the model is shown in Figure 1b. An example of the use of this model with respect to pavements was given by Harr (3).

The concept of a stochastic system is shown in Figure 2, which illustrates a vertical cross section through a pavement subjected at its surface (the x-axis) to a unit force (say per unit length normal to the plane of the paper) acting at point $x = x_1$. Suppose that we wish to determine the magnitudes of the two forces F_A and F_B, located as shown at equal distances a on either side of the unit force at a constant depth $z = z_1$. In effect, we would seek the transmission of the unit force through the pavement. A pavement, in its general form, is composed of a complex conglomeration of discrete particles, in arrays of varying shapes, sizes, and orientations, and contains randomly distributed concentrations of cementing agents. Certainly, we cannot expect that, in general, the forces registered by F_A and F_B will always be equal. (In the deterministic approach, it is customary to plead symmetry and hence the equality of the two forces, i.e., $F_A = F_B$.) In fact, because of the variations in the characteristics of such media, we should expect that they will seldom be equal because the location $x = x_1$ varies. For example, if one of the forces was not in contact with a solid particle, i.e., was in a void, no force would be noted. Evidently, as the unit force is moved in time, as for moving vehicles, through a series of points $x = x_1$, the magnitudes of the forces would be expected to be random in character, i.e., casual rather than causal. Systems that display random results with time are said to be stochastic. The thesis here is that, to be meaningful, experiments involving such systems should be formulated in terms of probabilistic statements rather than explicit expressions.

RANDOMNESS AND PROBABILITY

As noted, in the deterministic approach the outcomes of experiments (observations or phenomena) are treated as absolute quantities. For example, when given the total weight W, total volume V, and weights of solids W_s for a water-saturated soil mass, the porosity can be obtained from the formula $n = (W - W_s)/V\gamma_w$. In particular, given $W = 100$ g, $W_s = 55$ g, $V = 100$ cm^3, the porosity is $n = 45$ percent. Implied in this result is that, if we were to carry out the weighings and volumetric determinations on a large number of samples of the soil, under certain similar conditions, we would expect on the average that the ratio of volume of voids to total volume would be 0.45. Obviously, the determined porosity would not be 45 percent in every experiment. Sometimes it would be 40 or 41 percent, other times 43 or 46 percent. Occasionally, it may even be very much smaller or very much greater than 45 percent.

This example illustrates what is meant by random experiments, experiments that can give varying outcomes (results) depending on chance circumstances that are either unknown or beyond control. The distributions of particle sizes in a number of soil samples from the same test pit will not be the same. The variation in measured pavement thickness for a given section will show considerable differences. Contrary to common belief, the inability to obtain concise descriptions of events or observations is not a declaration of ignorance; it is the way nature and the real world behave—fraught with uncertainty. Stated more succinctly, there is no absolute knowledge. What physicists considered exact and ordered prior to the development of quantum mechanics turned out to be merely the mean value of a much more impressive structure.

The diversity of results of apparently similar circumstances is the consequence of randomness. A single data set represents only one of many possible results that may occur. Each of these may be considered as a single result of a random experiment (or phenomenon), the collection of which produces a random process. In other words, a data record for a particular sample of a random phenomenon is only one physical realization of a random process.

Every test and measurement conducted on pavement systems introduces a magnitude whose numerical value depends on random factors that are beyond control. Furthermore, each such magnitude can have a different value in successive trials. This type of magnitude is called a random variable; the separate magnitudes are called elementary events. The outcomes of a random experiment are called elementary events if (a) only one outcome can occur at a time and (b) one outcome always does occur. Condition a specifies that the outcomes are mutually exclusive or are disjoint; that is, no two elementary events can occur simultaneously. Condition b states that an elementary event is possible. The classic example of an elementary event is the outcome of the toss of a fair die. (Historically, questions relative to dice, asked of Pascal, precipitated the mathematical theory of probability.) Condition a is satisfied because only one face can appear per toss of a die. Condition b is ensured because any one of the six faces is likely to appear.

Implied in the die toss experiment is that the numbers 1, 2, 3, 4, 5, and 6 each have a possibility of occurring with equal likelihood. However, the number that will appear on any one toss is uncertain. Suppose now that the experiment is repeated many times. Even though the numbers shown on the faces may be different in successive tosses, it is reasonable to expect that, over the long run, any one number will occur one-sixth of the time. A gambler would say that the odds against tossing any specified number is five to one. The probabilist would define that probability to be one-sixth.

The measure of the probability of an outcome is its relative frequency. That is, if an outcome E can occur n times in N equally likely trials, the probability of the occurrence of outcome E (after a large, theoretically infinitely large, number of experiments) is

$$P(E) = \frac{n}{N} \tag{2}$$

Also, implied in equation 2 is the concept of the ratio of favorable outcome to the number of all possible cases. This definition was first formulated by Laplace in 1812. Stated another way, the probability of outcome A equals the number of outcomes favorable to A divided by the total number of outcomes or

$$P(A) = \frac{\text{favorable outcomes}}{\text{total outcomes}} \tag{3}$$

As an example, find the probability of drawing a red card from an ordinary well-shuffled deck of 52 cards. Of the 52 mutually exclusive and equally likely outcomes (each card is a possible outcome and, hence, an elementary event), there are 26 favorable outcomes (red cards); hence,

$$P(\text{drawing a red card}) = \frac{\text{favorable outcomes}}{\text{total outcomes}} = \frac{26}{52} = \frac{1}{2}$$

Knowledge of the distribution of a random variable and of the probabilities of the various possible values enables predictions to be made of the occurrence of an event or collection of events. The underlying question is, How does one find this distribution? Considering the uses to which probability theory has been put and the length of time it has been around, it is not surprising that there are many avenues available. At this point in our development, it will suffice to list a few examples.

1. Some few experiments are performed, and frequencies are noted and then generalized; e.g., a fair coin will show heads half of the time, or the porosity of the soil layer is 42 percent.
2. Probabilities are assigned subjectively as a set of weights that express likelihoods of outcomes; e.g., each of the 10 questions in the examination is worth 10 points, and 70 is the minimum passing grade.
3. Whole families of distribution laws stem from mathematical excursion derived from certain intuitive concepts. Detailed examples are available in texts dealing with probability (2).

Whatever the basis for assigning probabilities to elementary events or outcomes, the following are axiomatic: The probability of an outcome A ranges between zero and unity

$$0 \leq P(A) \leq 1 \tag{4}$$

and the certainty of an outcome C has a probability of unity

$$P(C) = 1 \tag{5}$$

The probability of the occurrence of a number of elementary events or outcomes A_1, $A_2, \ldots, A_i$ is the sum of component probabilities (addition rule):

$$P(A_1 + A_2 + \ldots + A_i) = P(A_1) + P(A_2) + \ldots + P(A_i) = \sum_{n=1}^{n=i} P(A_n) \tag{6}$$

This also implies that

$$P(A_1 \text{ or } A_2 \text{ or } \ldots A_i) = P(A_1) + P(A_2) + \ldots + P(A_i) \tag{7}$$

Equations 4, 5, 6, and 7 are the building blocks from which the elementary theory of probability is constructed. These axioms will next be used to derive some important properties of the probabilities of mutually exclusive outcomes.

A collection of elementary outcomes, $A_1, A_2, \ldots, A_n$, is said to be collectively exhaustive if it represents all outcomes that can occur for a given experiment. Symbolically, this is shown as

$$S = \{A_1, A_2, \ldots, A_n\} \tag{8}$$

For the toss of a die the six outcomes—$A_1 = 1$ appears, $A_2 = 2$ appears, ..., $A_6 = 6$ appears—are both mutually exclusive and collectively exhaustive; therefore,

$$S = \{A_1, A_2, \ldots, A_6\} \tag{9}$$

From equations 5 and 6 it follows that the probability of a collectively exhaustive set of outcomes is unity:

$$P(S) = P(A_1) + P(A_2) + \ldots + P(A_n) = 1 \tag{10}$$

An event A^c is said to be the complement of outcome A if $A^c + A = 1$. Here $S = \{A, A^c\}$. From equation 10 it follows that

$$P(A^c) = 1 - P(A) \tag{11}$$

It follows from equations 5 and 11 that, if C is the certain outcome, its complement is the uncertain outcome U whose probability is

$$P(U) = 0 \tag{12}$$

As another example, let w represent the square opening of a particular sieve of wire diameter D, and find the probability that a spherical particle of diameter d will hit the wire if the particle falls perpendicular to the plane of the sieve (Figure 3).

The possible outcomes can be related to the location of the center of the sphere relative to the mesh. It will only be necessary to consider one square because all squares present similar situations. The probability of the sphere hitting the mesh can be measured by the likelihood (favorable outcome) of the center of the sphere hitting beyond the dotted square ABCE in Figure 3. The total outcome is measured by the area of the square abce = $(w + D)^2$. Thus, the favorable outcome is the area between square ABCE and square abce = $(w + D)^2 - (w - D - d)^2$. Hence, the probabilities are

$$P(\text{hitting mesh}) = 1 - \left(\frac{w - D - d}{w + D}\right)^2 \tag{12a}$$

and

$$P(\text{passing through opening}) = \left(\frac{w - D - d}{w + D}\right)^2 \tag{12b}$$

This example provides some interesting excursions into what happens during the sieving process. If we assume that there is no interference between particles and that for a No. 200 sieve w = 0.074 mm and D = 0.0021 mm, a particle with a diameter 90 percent of the opening d = 0.067 mm has a probability of 0.9959 of hitting the mesh under the stated conditions. Suppose now that the sieve is shaken through n cycles. If we assume each cycle is equivalent to repeating the sphere dropping process, the probability of the particle hitting the mesh after n cycles is $(0.9959)^n$. To achieve a probability of 90 percent that the particle will pass through the opening—a probability of 0.10 that it will hit the mesh—will require $(0.9959)^n = 0.10$ or n = 554 cycles, hardly a manual task!

Now consider the cross section through a sample of a particulate medium at elevation z (Figure 4). The shaded portion defines the intersection of the plane of the cross section with the solid particles; the unshaded portion represents the region of voids. If m(z) is the area porosity, the ratio of the area of voids $A_v(z)$ (shown unshaded) to the total area A of the section at any elevation z, then

$$m(z) = \frac{A_v(z)}{A}$$

The average value of m(z) over the height of the sample h is

$$m = \frac{1}{h}\int_0^h m(z)dz$$

Then

$$m = \frac{1}{Ah}\int_0^h Am(z)dz = \frac{1}{V}\int_0^h A_v(z)dz$$

where V is the total volume and the integral is the volume of voids. Hence, the average value of m is the volume porosity n. Based on the geometric definition of probability, it follows that the porosity of a particulate material $n = V_v/V$ is the probability of finding (hitting) a void in a unit volume. The complementary event to the porosity is the volume of solids per unit volume $n_s = 1 - n$; therefore, the probability of locating a solid particle in a unit volume of material is

$$P(\text{locating a particle}) = n_s = 1 - n \tag{13}$$

In the limit, when there are no voids, equation 13 shows that the certainty of locating a solid is $n_s = 1$.

TREE DIAGRAMS

The characterization of a sample of a highway pavement requires a choice of one of three identification tests (I_1, I_2, I_3), a choice of one of three strength tests (S_1, S_2, S_3), and a choice of one of two compressibility tests (C_1, C_2). In how many possible ways may the characterization of tests be performed? The three-step process is shown in Figure 5. It should be evident why this representation is called a tree diagram. The total number of outcomes is the same as number of possible paths, which in this case is simply the sum of the twigs on the lowest branch, i.e., the number of C's, which is 18. The ordering of the tests is arbitrary.

As another example, consider that there are four classifications of soils (gravel, sand, silt, and clay), the soils may be saturated or unsaturated, and they may exhibit low, medium, and high degrees of density. If all sequences are equally likely, how probable is it that a randomly selected sample will be a dense, saturated clay?

The tree diagram is shown in Figure 6. There are 24 possibilities; hence the probability of selecting a sample of dense, saturated clay (shown darker) is 1/24.

It should be apparent that the tree diagram is in a sense a graphical multiplier. The same results could have been obtained by using a multiplication rule. Stated simply, if the first step (or event) A_1 has a_1 outcomes and for each of these a_1 outcomes a second independent step (or event) A_2 has a_2 outcomes, then there will be the product a_1a_2 outcomes after the two steps. Obviously, this can be extended to any number of independent events. For the example in Figure 5, this would lead to $3 \times 3 \times 2 = 18$ outcomes. For Figure 6, $4 \times 2 \times 3 = 24$ outcomes.

PERMUTATIONS AND COMBINATIONS

The listing of all the distinct arrangements of r objects within a collection of n objects is called the permutations of n objects taken r at a time. Symbolically, this is shown as

$$P(n, r) = \frac{n!}{(n - r)!} \tag{14}$$

where $n! = n(n - 1)(n - 2) \ldots (2)(1)$. Other common designations are ${}_nP_r$, P_r^n, and $P_{n,r}$.

Arranging n objects in some order of r is the same as preparing a tree diagram with n ways shown in the first step, n - 1 ways in the second, until n - r + 1 ways occupy the last (the r th) position.

If there are 10 automobiles that at various times occupy the six available parking spaces reserved for employees at the rear of a store, how many different parking arrangements are possible assuming that no parking spaces remain unoccupied?

The question asks, How many groups of six can be arranged from among 10 objects? Thus, n = 10, r = 6, and $P(10, 6) = 10 \times 9 \times 8 \times 7 \times 6 \times 5 = 151{,}200$.

Another common situation is when one wishes to find the number of distinct permutations (remembering that all permutations must be different) when some of the objects in the collection are alike. Suppose that of the n objects n_1 are all of one kind, n_2 are all of a second kind, etc., and there are k kinds. It can be shown that the number of distinct permutations is

$$\frac{n!}{(n_1!\, n_2! \ldots n_k!)} \tag{15}$$

In how many ways can 15 objects be divided into two groups such that one has twice as many as the other? Here n = 15, $n_1 = 10$, and $n_2 = 5$. Hence applying equation 15 gives $15!/(10! \cdot 5!) = 3{,}003$ ways.

A combination of objects rather than a permutation occurs when order is of no importance. Generally, a combination can be thought of as a selection and a permutation as an arrangement. A combination of n objects taken r at a time is a selection of r objects taken from among the n objects without specifying order or arrangement. For example, the combination of the four letters a, b, c, and d taken three at a time is abc, abd, acd, and bcd. Because order is not important, the combinations abc, acb, cab, cba, bac, and bca are all equal. The number of combinations of r objects from among n objects is given as

$$C(n, r) = \frac{n!}{r!(n - r)!} \tag{16}$$

This sequence follows from the observation that there are r! permutations of every combination of $P(n, r) = r!\,C(n, r)$. It follows from equation 16 that $C(n, r) = C(n, n - r)$. Other designations are ${}_nC_r$, $C_{n,r}$, C_r^n. Another commonly used symbol for the r combinations of n objects is

$$C(n, r) = \binom{n}{r} \tag{17}$$

An inspector on a highway project discovers that k of every n samples tested can be expected to be below standard. If n samples are chosen randomly, from a batch of n samples find the probability that ℓ are below standard.

There are a total of $\binom{n}{m}$ number of ways to choose m items out of n. The number of favorable ways is $\binom{k}{\ell} \times \binom{n-k}{m-\ell}$; that is, the favorable outcomes (for this example, not for the inspector) are those in which ℓ substandard tests are found among k items, which can be done $\binom{k}{\ell}$ ways, and the remaining m - ℓ tests are at or above standard from among the total number of n - k tests. Hence the required probability will be the number of favorable ways divided by the total number of ways or

$$P = \frac{\binom{k}{\ell} \times \binom{n-k}{m-\ell}}{\binom{n}{m}} \tag{18}$$

CONDITIONAL PROBABILITIES

Concrete for a particular highway pavement is mixed in two plants P_1 and P_2 and trucked to the jobsite. Plant P_1 produces 66 percent and plant P_2 34 percent of the concrete used. Tests on concrete cylinders judge the concrete to meet standards if the 28-day unconfined compressive strength is no less than 4,500 psi (31 000 kPa). Previous tests on the concrete from plant P_1 show that an average of 91 percent of the concrete meets standards. Results from plant P_2 show that 83 percent of the samples meet the criterion. These results would indicate that the highway inspector can expect, on the average, that approximately 88 tests per 100 will prove adequate $[(0.66) \times (91) + (0.34) \times (83) \approx 88]$. Stated another way, the probability of a sample of concrete meeting the specification is approximately 0.88. If all the concrete was obtained from plant P_1, the probability of getting standard concrete would be 91 percent; plant P_2 would produce 83 percent. It is apparent that information on where the concrete was produced will affect the probability of the outcome. Such probabilities are said to be conditional; that is, the occurrence of one outcome (information on which plant produced the concrete) will modify the chance of the occurrence of another outcome (the probability that a number of test samples of the concrete will test above standard). Before the origin of the concrete was specified, the unconditional probability that the tests would pass the specification was 0.88.

The conditional probability of an outcome A, given that an outcome B has occurred, denoted $P(A|B)$, is defined as

$$P(A|B) = \frac{P(AB)}{P(B)} \tag{19}$$

where $P(AB)$ denotes the probability that both (simultaneous occurrences) outcomes A and B will occur, called their intersection, and $P(B)$ is the probability of the occurrence of outcome B. (In set theory AB, designated $A \cap B$, is the intersection of A and B or the joint occurrence of A and B.) If $P(B) = 0$, the conditional probability is not defined. Other useful forms of this expression are

$$P(AB) = P(A)\, P(B|A)$$

or

$$P(AB) = P(B)\,P(A\,|\,B) \tag{20}$$

These forms yield the probability of the intersection (or simultaneous occurrences) of both outcomes A and B. In words, the simultaneous occurrence of two outcomes is equal to the product of the probability of one outcome and the conditional probability of the other, assuming the first occurred.

For example, a pair of fair dice are thrown. What is the (conditional) probability that their sum is greater than 6 if a 2 appears on the first die?

There are $6^2 = 36$ possible outcomes. The number 2 appearing on the first die has a probability of occurrence of $1/6$. There are two (simultaneous) favorable outcomes: a five or six on the second die. This is the outcome AB. Hence, $P(AB) = 2/36$. Using equation 19 gives a conditional probability of

$$P(A\,|\,B) = \frac{2/36}{1/6} = 2/6 = 1/3$$

An example of the unconditional probability for this example would correspond to asking, What is the probability that a sum greater than 6 will appear on the throw of a pair of dice? Because there are 21 favorable outcomes, the unconditional probability would equal $21/36 = 0.58$.

For a number of problems a tree diagram is very useful in understanding the outcomes and their probabilities.

A box contains 10 articles; six are painted red and four are white. If two articles are selected at random (without replacement), what are the probabilities of the various permutations?

The results are shown on the tree diagram in Figure 7. The probability of any permutation (path) is the product of the branch probabilities. On the first draw, the probability of drawing red is $P(R) = 6/10$. However, on the second draw, because there are only 5 reds available among the 9 articles, $P(R\,|\,R) = 5/9$. Hence from $P(RR) = P(R) \times P(R\,|\,R) = 1/3$. Note that the probability of drawing one of each color at the end of the drawing is $P(RW) + P(WR) = 4/15 + 4/15 = 8/15$.

Equation 20 gives the probability of the joint occurrence (intersection) of two outcomes A and B. These can be further generalized to any number of outcomes. For three outcomes, ABC,

$$P(ABC) = P(A) \times P(B\,|A) \times P(C\,|\,AB) \tag{21}$$

where $P(C\,|\,AB)$ is the probability that outcome C occurs given that the joint outcomes of A and B have already occurred.

Find the probability of drawing hearts on three consecutive draws, without replacement, from a standard deck of cards. We define A, B, C as outcomes of a heart being drawn on the first, second, and third draws. Here as in the previous example $P(A) = 1/4$, $P(B\,|\,A) = 4/17$, and $P(C\,|\,AB) = 11/50$. Hence,

$$P(\text{drawing three consecutive hearts}) = 1/4 \times 4/17 \times 11/50 = 0.013$$

Equation 21 can be extended to any number of outcomes. For any outcomes $A_1, A_2, \ldots, A_n$,

$$P(A_1A_2\ldots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\ldots P(A_n|A_1A_2\ldots A_{n-1})$$

or

$$P\left(\prod_{k=1}^{n} A_k\right) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\ldots P\left(A_n \middle| \prod_{k=1}^{n-1} A_k\right) \tag{22}$$

Equation 22 is called the multiplication rule of probability; the addition rule was given in equation 6.

A sample of gravel from a gravel pit is to be examined to determine whether the pit can furnish adequate coarse aggregate for making concrete. Specifications require that the pit be rejected if at least one deleterious particle is discovered among five particles selected at random from the sample. What is the probability that the gravel pit will be acceptable if a sample of 100 particles contains 5 deleterious particles?

Let A_i be finding a nondeleterious particle as the i th outcome. The probability of acceptance is the probability of the joint outcome $P(A_1A_2A_3A_4A_5)$. The probability of not finding the first deleterious particle is $P(A_1) = 95/100$ because, of the 100 particles, 95 are acceptable. After the occurrence of event A_1, there remain 99 particles of which 94 are not deleterious; hence, $P(A_2|A_1) = 94/99$. Continuing the reasoning, $P(A_3|A_1A_2) = 93/98$, $P(A_4|A_1A_2A_3) = 92/97$, and $P(A_5|A_1A_2A_3A_4) = 91/96$. Therefore, from equation 22,

$$P(\text{acceptance}) = 95/100 \times 94/99 \times 93/98 \times 92/97 \times 91/96 = 0.77$$

The same result can be obtained another way. There are $\binom{95}{5}$ favorable ways that 5 good particles can be selected from among 95 good ones. There are a total of $\binom{100}{5}$ ways of selecting 5 particles from among 100. Hence,

$$P(\text{acceptance}) = \binom{95}{5} \Big/ \binom{100}{5} = 0.77$$

A useful graphical representation of various operations involving outcomes is the Venn diagram. The outcomes are usually represented by simple geometrical shapes. Suppose now that the rectangular region S shown in Figure 8 represents the total probability $P(S) = 1$. Then the probabilities associated with any outcomes are the sum of the elementary outcomes that are contained within their respective regions in the Venn diagram. For example, suppose the dots shown in the figure represent elementary outcomes; those belonging to A are within circle A and those for B within the circle. Where the two circles overlap (shown shaded), some elementary outcomes in A are also in B. This corresponds to the intersection of A and B, which is designated AB or $A \cap B$. The probability that A or B or P(A or B) (of the union of A and B, $A \cup B$) will occur is the sum of the probabilities of their respective elementary outcomes with each elementary outcome accounted for only once. Because the regions overlap, the simple sum of their elementary outcomes $P(A) + P(B)$ would add the elements in the common region twice; hence, in general,

$$P(A \text{ or } B) = P(A) + P(B) - P(AB) \tag{23a}$$

or in set notation

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad (23b)$$

For the example with the concrete plant mentioned previously, the following outcomes are defined:

A = concrete meets standard,
A^c = concrete does not meet standard,
B = concrete produced at first plant, and
B^c = concrete produced at second plant.

It is immediately apparent that there are no outcomes simultaneously favorable to $P(AA^c)$ and $P(BB^c)$. Such outcomes are mutually exclusive or incompatible; both cannot happen simultaneously. On the other hand, some concrete produced at both plants did not meet standards; hence,

$$P(A) = P(AB) + P(AB^c) \qquad (24a)$$

$$P(B) = P(AB) + P(A^cB) \qquad (24b)$$

What is the probability of the outcomes A or B or both or P(A or B) = ? In other words, what is the probability that the concrete meets the standard or is produced at the first plant or that both conditions are satisfied? From equation 6, it follows that

$$P(A \text{ or } B) = P(AB) + P(AB^c) + P(A^cB)$$

All other possibilities are incompatible. Combining equations 24a and 24b gives the same result as in equation 23.

Of a sample of 100 particles inspected, 15 are deleterious, 8 are too large for the intended use, and 6 have both defects. If a particle is selected at random from the sample, what is the probability that it will not be suitable?

$$P(\text{not suitable}) = P(\text{deleterious}) + P(\text{too large}) - P(\text{both})$$

$$= {}^{15}/_{100} + {}^{8}/_{100} - {}^{6}/_{100} = 0.17$$

Equation 23 demonstrates that, if $P(AB) = 0$, the probability of the joint occurrence is zero and equation 6 applies. The outcomes are said to be independent. Stated another way, two outcomes are independent if the occurrence or nonoccurrence of one has no effect on the probability of the occurrence of the other. Independence implies unconditional probability or $P(A|B) = P(A)$. Hence, for the case of the independence of outcomes, the addition rule is

$$P(A_1 + A_2 + A_3 + \ldots + A_i) = P(A_1) + P(A_2) + \ldots + P(A_i) \qquad (25)$$

and the multiplication rule is

$$P(A_1A_2A_3 \ldots A_i) = P(A_1) \times P(A_2) \times \ldots \times P(A_i) \quad (26)$$

Find the probability of drawing a heart on three consecutive draws from a standard deck if the cards are replaced and the deck is reshuffled after each draw. As stated, each draw is independent of the other; hence, the probability of drawing a heart each time is $1/4$, and

$$P(\text{drawing three consecutive hearts}) = 1/4 \times 1/4 \times 1/4 = 0.016$$

Without replacement, the probability is slightly less, 0.013.

The scheme shown in Figure 9 represents the road system between two cities. R_M and R_N are city roads, and S_1, S_2, and S_3 are state highways. After a major snowstorm the probabilities of the various highways being impassable are as shown in the figure. What is the probability of a driver being able to get from city M to city N under the stated conditions if his or her choice of state highway is random?

The probability of the system being open is the complement of it being closed. If p is the probability of impassability,

$$p = P(R \text{ or } S) = P(R) + P(S) - P(RS)$$

where R and S are the outcomes that a city highway or a state highway is impassable. Hence, $P(R) = P(R_M) + P(R_N) - P(R_M)\,P(R_N)$ because outcomes R_M and R_N are independent. That is, $P(R_MR_N) = P(R_M) \times P(R_N|R_M) = P(R_M) \times P(R_N)$. Thus, $P(R) = 0.6 + 0.5 - (0.6) \times (0.5) = 0.8$. $P(S) = P(S_1)\,P(S_2)\,P(S_3)$ because each state highway is independent of the other. Thus $P(S) = (0.4)(0.7)(0.9) = 0.252$. Furthermore, because R and S are independent, $P(RS) = P(R) \times P(S) = (0.8)(0.252) = 0.202$. Therefore, $p = 0.800 + 0.252 - 0.202 = 0.85$ and $P(\text{driving between the cities}) = 1 - 0.85 = 0.15$.

BAYES' THEOREM

Suppose that n outcomes $A_1A_2A_3, \ldots, A_n$ are mutually exclusive and collectively exhaustive such that $\sum_{i=1}^{n} P(A_i) = 1$. Suppose also that there is another outcome B whose occurrence was preceded by or caused by one of the n outcomes of A_i; which of the outcomes in A_i is not known. Because an outcome in A_i precipitated B, it is said to be prior or an a priori outcome. The occurrence of outcome B (called an a posteriori outcome) requires the occurrence of one of the joint outcomes A_iB, which, in general, can be written as

$$P(B) = \sum_{i=1}^{n} P(A_iB)$$

Using the multiplication rule gives

$$P(B) = P(B|A_1) \times P(A_1) + P(B|A_2) \times P(A_2) + \ldots + P(B|A_n) \times P(A_n)$$

$$= \sum_{i=1}^{n} P(B|A_i)P(A_i) \qquad (27)$$

Equation 27 is called the total probability theorem.

A boring record at the site of a bridge abutment for an overpass indicates that 30 percent of the soil profile of interest is sand, 25 percent silty-sand, 25 percent silty clay, and 20 percent clay. Samples are taken at the site in proportion to layer thickness; however, for various reasons, not all the samples are reliable. Indications are that only 27 percent of the sand samples are adequate, 10 percent of the silty-sand, 30 percent of the silty clay, and 50 percent of the clay. If we wish to examine in detail the soil characteristics at a particular depth, what is the probability that one of the samples chosen at random will furnish reliable information?

Designate the initial samples as A_s, A_{ss}, A_{sc}, and A_c. The prior probabilities were $P(A_s) = 0.3$, $P(A_{ss}) = 0.25$, $P(A_{sc}) = 0.25$, $P(A_c) = 0.20$. If B denotes that the sample at the desired depth is among the reliable samples, the probable occurrences are $P(B|A_s) = 0.27$, $P(B|A_{ss}) = 0.10$, $P(B|A_{sc}) = 0.30$, $P(B|A_c) = 0.50$. Hence, the required probability is

$$P(B) = (0.3)(0.27) + (0.25)(0.10) + (0.25)(0.30) + (0.20)(0.50) = 0.28$$

Again assume that the prior outcomes A_1, $A_2, \ldots, A_n$ are mutually exclusive and collectively exhaustive. Also suppose that the outcome of B is preceded by or caused by one of the A_i. Again, which one is not known. From equation 20,

$$P(A_iB) = P(A_i) \times P(B|A_i)$$

or

$$P(A_iB) = P(B) \times P(B|A_i)$$

from which

$$P(A_i|B) = \frac{P(A_i)\,P(B|A_i)}{P(B)} \qquad (28)$$

Expressing P(B) by using the total probability equation (equation 27) produces

$$P(A_i|B) = \frac{P(A_i) \times P(B|A_i)}{\sum_{i=1}^{n} P(B|A_i) \times P(A_i)} \qquad (29)$$

This result is called Bayes' theorem. It indicates how opinions held before an experiment should be modified by the evidence of the outcome. Historically, its statement

Figure 1. (a) Deterministic system and (b) model.

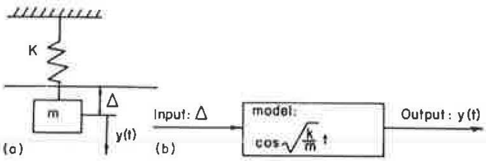


Figure 2. Example of stochastic system.

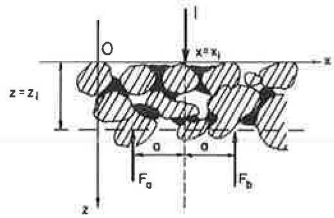


Figure 3. Sieve opening example.

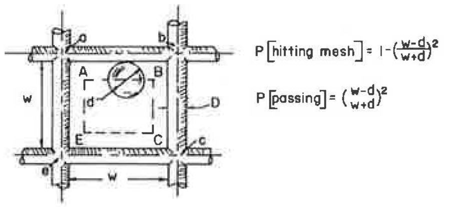


Figure 4. Example for locating particle in sample.

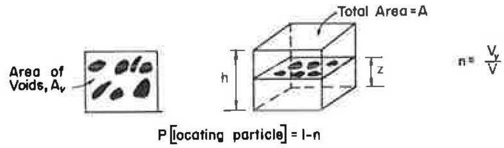


Figure 5. Three-step tree diagram.

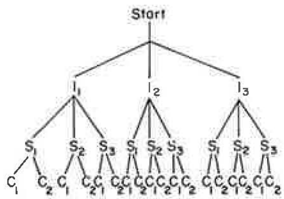


Figure 6. Probability of selecting favorable alternative = $^1/_{24}$.

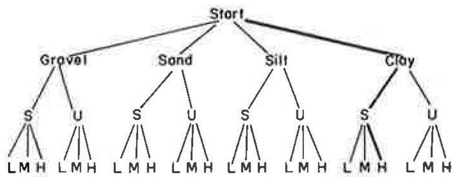


Figure 7. Probabilities of permutations.

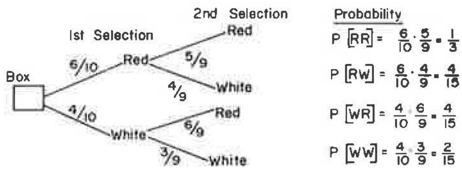
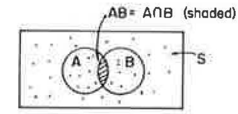


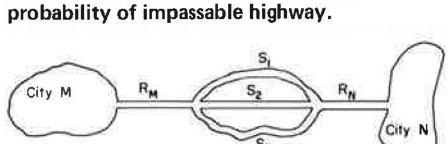
Figure 8. Venn diagram.



P[A or B] = P[A] + P[B] - P[AB]

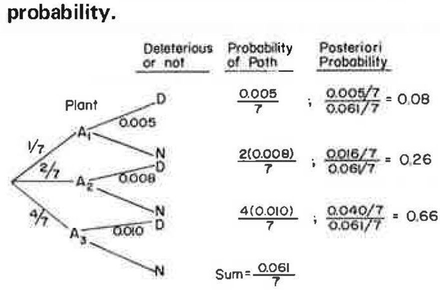
P[A ∪ B] = P[A] + P[B] - P[A∩B]

Figure 9. Scheme for determining the probability of impassable highway.



Highway	P[impossible]
R_M	0.6
R_N	0.5
S_1	0.4
S_2	0.7
S_3	0.9

Figure 10. Tree diagram of a posteriori probability.



marked the beginning of the mathematical theory of inductive reasoning.

When written in the following form, equation 29 demonstrates how the introduction of the a posteriori outcome B alters the a priori assessment of the probability of A_i:

$$P(A_i\,|\,B) = P(A_i)\ \{\text{modification of } A_i \text{ when B is learned of}\} \qquad (30)$$

where A_i are prior probabilities.

Aggregates used for highway construction are produced at three plants with daily production volumes of 500, 1,000, and 2,000 tons (450, 900, 1800 Mg). Past experience indicates that the fractions of deleterious materials produced at the three plants are 0.005, 0.008, and 0.010 respectively. If a sample of aggregate is selected at random from a day's total production and found to be deleterious, which plant is it likely produced the sample? Designate the following outcomes:

A_1 = production from the first plant, 500 tons (450 Mg) per day.
A_2 = production from the second plant, 1,000 tons (900 Mg) per day.
A_3 = production from the third plant, 2,000 tons (1800 Mg) per day.

The prior probabilities are

$$P(A_1) = \frac{500}{500 + 1{,}000 + 2{,}000} = 1/7$$

$$P(A_2) = \frac{1{,}000}{3{,}500} = 2/7$$

$$P(A_3) = \frac{2{,}000}{3{,}500} = 4/7$$

The likelihoods are

$$P(B\,|\,A_1) = 0.005$$

$$P(B\,|\,A_2) = 0.008$$

$$P(B\,|\,A_3) = 0.010$$

The joint occurrences are

$$P(A_1)\,P(B\,|\,A_1) = 1/7\,(0.005)$$

$$P(A_2)\,P(B\,|\,A_2) = 2/7\,(0.008)$$

$$P(A_3)\,P(B\,|\,A_3) = 4/7\,(0.010)$$

$$\text{Sum} = \Sigma P(A_i)\,P(B|A_i) = \frac{0.061}{7}$$

A posteriori probabilities are

$$P(A_1|B) = \frac{0.005/7}{0.061/7} = {}^{5}/_{61} = 0.08 \quad (0.14 \text{ a priori})$$

$$P(A_2|B) = \frac{0.016/7}{0.061/7} = {}^{16}/_{61} = 0.26 \quad (0.29 \text{ a priori})$$

$$P(A_3|B) = \frac{0.040/7}{0.061/7} = {}^{40}/_{61} = 0.66 \quad (0.57 \text{ a priori})$$

Because $P(A_3|B) = 0.66$ presents by far the greatest a posteriori probability, it is most probable that the deleterious sample came from the third plant. Of course, the same conclusion would be valid from a priori probabilities in this case. Figure 10 shows the tree diagram for this example. The a posteriori probabilities are obtained as the ratio of the probability of the required path to that of the sum of all paths that lead to a particular deleterious sample.

REFERENCES

1. J. R. Benjamin and C. A. Cornell. Probability, Statistics, and Decision for Civil Engineers. McGraw-Hill, 1970.
2. W. Feller. An Introduction to Probability Theory and Its Applications. John Wiley and Sons, 1957.
3. M. E. Harr. Influence of Vehicle Speed on Pavement Deflections. HRB Proc., Vol. 41, 1962.
4. M. E. Harr. Particulate Mechanics. In press.
5. T. C. Fry. Probability and Its Engineering Uses. D. Van Nostrand Co., 1965.

DISCUSSION

Richard L. Davis, Koppers Company, Pittsburgh

Traditional engineering education has had most calculations starting with assumptions without very much guidance on how to make valid assumptions. In nearly all engineering, these assumptions have grossly oversimplified the real conditions. This is particularly true of highway engineering where there is a great need to deal with real-world phenomena. I believe that Harr's paper can make an important contribution to this process for engineers. This is no easy task for the average engineer who usually views a test result for asphalt content of a mix as its true value rather than as a probability estimate of the true asphalt content. Familiarity with and use of probability methods will help to clarify and solve many baffling problems in highway engineering.