



IGNITED MINDS
Journals

*Journal of Advances in
Science and Technology*

*Vol. 11, Issue No. 22,
May-2016, ISSN 2230-9659*

**AN ANALYSIS UPON CONCEPT AND BASICS OF
PROBABILITY THEORY: A REVIEW**

AN
INTERNATIONALLY
INDEXED PEER
REVIEWED &
REFEREED JOURNAL

An Analysis upon Concept and Basics of Probability Theory: A Review

Rajkumar Ahuja^{1*} Dr. Vinod Kumar Sharma²

¹Professor

²Professor, Department of Mathematics, Faculty of Science, Tanta University, Sri Ganganagar, Rajasthan, India

Abstract – In recent years probabilistic knowledge-based systems such as Bayesian networks and influence diagrams have come to the fore as a means of representing and reasoning about complex real-world situations. Although some of the probabilities used in these models may be obtained statistically, where this is impossible or simply inconvenient, modellers rely on expert knowledge. Experts, however, typically find it difficult to specify exact probabilities and conventional representations cannot reflect any uncertainty they may have. In this way, the use of conventional point probabilities can damage the accuracy, robustness and interpretability of acquired models.

In talking about no discrete probability spaces, it is difficult to avoid measure-theoretic concepts. However, to develop extensive formal machinery from measure theory before going into probability (as is done in most graduate programs in mathematics) would be inappropriate for the particular audience to whom the book is addressed. Thus I have tried to suggest, when possible, the underlying measure-theoretic ideas, while emphasizing the probabilistic way of thinking, which is likely to be quite novel to anyone studying this subject for the first time.

INTRODUCTION

The origin of probability theory lies in physical observations associated with games of chance. It was found that if an “unbiased” coin is tossed independently n times, where n is very large, the relative frequency of heads, that is, the ratio of the number of heads to the total number of tosses, is very likely to be very close to $1/2$. Similarly, if a card is drawn from a perfectly shuffled deck and then is replaced, the deck is reshuffled, and the process is repeated over and over again, there is (in some sense) convergence of the relative frequency of spades to $1/4$.

In the card experiment there are 52 possible outcomes when a single card is drawn. There is no reason to favor one outcome over another (the principle of “insufficient reason” or of “least astonishment”), and so the early workers in probability took as the probability of obtaining a spade the number of favorable outcomes divided by the total number of outcomes, that is, $13/52$ or $1/4$.

This so-called “classical definition” of probability (the probability of an event is the number of outcomes favorable to the event, divided by the total number of outcomes, where all outcomes are equally likely) is first of all restrictive (it considers only experiments with a finite number of outcomes) and, more seriously,

circular (no matter how you look at it, “equally likely” essentially means “equally probable,” and thus we are using the concept of probability to define probability itself). Thus we cannot use this idea as the basis of a mathematical theory of probability; however, the early probabilists were not prevented from deriving many valid and useful results.

Similarly, an attempt at a frequency definition of probability will cause trouble. If S_n is the number of occurrences of an event in n independent performances of an experiment, we expect physically that the relative frequency S_n/n should converge to a limit; however, we cannot assert that the limit exists in a mathematical sense. In the case of the tossing of an unbiased coin, we expect that $S_n/n \rightarrow 1/2$, but a conceivable outcome of the process is that the coin will keep coming up heads forever. In other words it is possible that $S_n/n \rightarrow 1$, or that $S_n/n \rightarrow$ any number between 0 and 1, or that S_n/n has no limit at all.

In this study we introduce the concepts that are to be used in the construction of a mathematical theory of probability. The first ingredient we need is a set Ω , called the *sample space*, representing the collection of possible outcomes of a random experiment. For example, if a coin is tossed once we may

take $\Omega = \{H, T\}$, where H corresponds to a head and T to a tail. If the coin is tossed twice, this is a different experiment and we need a different Ω , say $\{HH, HT, TH, TT\}$; in this case one performance of the experiment corresponds to two tosses of the coin.

If a single die is tossed, we may take Ω to consist of six points, say $\Omega = \{1, 2, \dots, 6\}$. However, another possible sample space consists of two points, corresponding to the outcomes “ N is even” and “ N is odd,” where N is the result of the toss. Thus different sample spaces can be associated with the same experiment. The nature of the particular problem under consideration will dictate which sample space is to be used. If we are interested, for example, in whether or not $N \geq 3$ in a given performance of the experiment, the second sample space, corresponding to “ N even” and “ N odd,” will not be useful to us.

In general, the only physical requirement on Ω is that a given performance of the experiment must produce a result corresponding to exactly one of the points of Ω . We have as yet no mathematical requirements on Ω ; it is simply a set of points.

Next we come to the notion of *event*. An “event” associated with a random experiment corresponds to a question about the experiment that has a yes or no answer, and this in turn is associated with a subset of the sample space. For example, if a coin is tossed twice and $\Omega = \{HH, HT, TH, TT\}$, “the number of heads is ≤ 1 ” will be a condition that either occurs or does not occur in a given performance of the experiment. That is, after the experiment is performed, the question “Is the number of heads ≤ 1 ?” can be answered yes or no. The subset of Ω corresponding to a “yes” answer is $A = \{HT, TH, TT\}$; that is, if the outcome of the experiment is $HT, TH, \text{ or } TT$, the answer

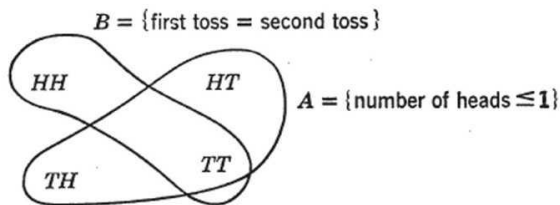


Figure 2.1 Coin-Tossing Experiment.

to the question “Is the number of heads ≤ 1 ?” will be “yes,” and if the outcome is HH , the answer will be “no.” Similarly, the subset of Ω associated with the “event” that the result of the first toss is the same as the result of the second toss is $B = \{HH, TT\}$.

Thus an *event* is defined as a subset of the sample space, that is, a collection of points of the sample space.

Events will be denoted by capital letters at the beginning of the English alphabet, such as A, B, C , and so on. An event may be characterized by listing all of its points, or equivalently by describing the conditions under which the event will occur. For example, in the coin-tossing experiment just considered, we write

$$A = \{\text{the number of heads is less than or equal to 1}\}$$

This expression is to be read as “ A is the set consisting of those outcomes which satisfy the condition that the number of heads is less than or equal to 1,” or, more simply, “ A is the event that the number of heads is less than or equal to 1.” The event A consists of the points $HT, TH, \text{ and } TT$; therefore we write $A = \{HT, TH, TT\}$, which is to be read “ A is the event consisting of the points $HT, TH, \text{ and } TT$.” As another example, if B is the event that the result of the first toss is the same as the result of the second toss, we may describe B by writing $B = \{\text{first toss} = \text{second toss}\}$ or, equivalently, $B = \{HH, TT\}$ (see Figure 2.1).

Each point belonging to an event A is said to be *favorable* to A . The event A will occur in a given performance of the experiment if and only if the outcome of the experiment corresponds to one of the

points of A . The entire sample space Ω is said to be the *sure* (or *certain*) event; it *must* occur on any given performance of the experiment. On the other hand, the event consisting of none of the points of the sample space, that is, the empty set \emptyset , is called the *impossible event*; it can *never* occur in a given performance of the experiment.

PROBABILITY

We now consider the assignment of probabilities to events. A technical complication arises here. It may

not always be possible to regard all subsets of Ω as events. We may discard or fail to measure some of the information in the outcome corresponding to the

point $\omega \in \Omega$, so that for a given subset A of Ω , it may not be possible to give a yes or no answer to the question “Is $\omega \in A$?” For example, if the experiment involves tossing a coin five times, we may record the results of only the first three tosses, so that $A = \{\text{at least four heads}\}$ will not be “measurable”; that is, membership of $\omega \in A$ cannot be determined from the given information about ω .

In a given problem there will be a particular class of subsets of Ω called the “class of events.” For

reasons of mathematical consistency, we require that the event class \mathcal{F} form a *sigma field*, which is a collection of subsets of Ω satisfying the following three requirements.

$$\Omega \in \mathcal{F} \quad (2.1)$$

$$A_1, A_2, \dots \in \mathcal{F} \text{ implies } \bigcup_{n=1}^{\infty} A_n \in \mathcal{F} \quad (2.2)$$

That is, \mathcal{F} is closed under finite or countable union.

$$A \in \mathcal{F} \text{ implies } A^c \in \mathcal{F} \quad (2.3)$$

That is, \mathcal{F} is closed under complementation.

Notice that if $A_1, A_2, \dots \in \mathcal{F}$, then $A_1^c, A_2^c, \dots \in \mathcal{F}$ by (2.3); hence $\bigcup_{n=1}^{\infty} A_n^c \in \mathcal{F}$ by (2.2). By the DeMorgan laws, $\bigcap_{n=1}^{\infty} A_n = (\bigcup_{n=1}^{\infty} A_n^c)^c$; hence, by (2.3), $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$. Thus \mathcal{F} is closed under finite or countable intersection. Also, by (2.1) and (2.3), the empty set \emptyset belongs to \mathcal{F} . Thus, for example, if the question "Did A_n occur?" has a definite answer for $n = 1, 2, \dots$, so do the questions "Did at least one of the A_n occur?" and "Did all the A_n occur?" Note also that if we apply the algebraic operations to sets in \mathcal{F} , the new sets we obtain still belong to \mathcal{F} .

In many cases we shall be able to take $\mathcal{F} =$ the collection of all subsets of Ω , so that every subset of Ω is an event. Problems in which \mathcal{F} cannot be chosen in this way generally arise in uncountably infinite sample spaces; for example, $\Omega =$ the reals.

We are now ready to talk about the assignment of probabilities to events. If $A \in \mathcal{F}$, the probability $P(A)$ should somehow reflect the long-run relative frequency of A in a large number of independent repetitions of the experiment. Thus $P(A)$ should be a number between 0 and 1, and $P(\Omega)$ should be 1.

Now if A and B are disjoint events, the number of occurrences of $A \cup B$ in n performances of the experiment is obtained by adding the number of occurrences of A to the number of occurrences of B . Thus we should have $P(A \cup B) = P(A) + P(B)$ if A and B are disjoint and, similarly, $P(A_1 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i)$ if

A_1, \dots, A_n are disjoint for mathematical convenience we require that

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

when we have a *countably infinite* family of disjoint events A_1, A_2, \dots .

The assumption of countable rather than simply finite additivity has not been convincingly justified physically or philosophically; however, it leads to a much richer mathematical theory.

A function that assigns a number $P(A)$ to each set A in the sigma field \mathcal{F} is called a *probability measure* on \mathcal{F} , provided that the following conditions are satisfied.

$$P(A) \geq 0 \text{ for every } A \in \mathcal{F} \quad (2.4)$$

$$P(\Omega) = 1 \quad (2.5)$$

If A_1, A_2, \dots are disjoint sets in \mathcal{F} , then

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots \quad (2.6)$$

We may now give the underlying mathematical framework for probability theory.

BASIC AXIOMS

The probability of an event is the area of the rectangle that represents the event, and the sample space is the union of all events. This representation can be generalized to more abstract spaces and leads to an axiomatic definition of probability in terms of measure over a collection of subsets. This collection is assumed to contain the empty set, and to be closed under the complementation and countable union (i.e. $\bigcup_{i=1}^{\infty} A_i \in \mathcal{S}$).

Theorem 2.1 Let S denote the sample space. A set

function $p(\cdot)$ defined in S is a probability function if:

1. For any event A in S , then $p(A) \geq 0$;

$$p(S) = 1;$$

3. If A_1, A_2, \dots are exclusive events in S and hence $A_i, A_j = \emptyset$ for all i, j , then $p(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} p(A_i)$

From these axioms, the following elementary properties can be derived.

Properties 1 Let $p(\cdot)$ be a probability function defined over the sample space S . Then $p(\cdot)$ satisfies the following properties:

1. $p(\emptyset) = 0$;

2. $p(\cdot)$ is finitely additive; if A_1, A_2, \dots, A_n are events in S , such that $A_i, A_j = \emptyset$ for all $i \neq j$, then

$$p(\cup_{h=1}^n A_h) = \sum_{h=1}^n p(A_h) \tag{2.7}$$

If these events form a partition of S , i.e. they are such that $\cup_{h=1}^n A_h = S$, then $p(\cup_{h=1}^n A_h) = 1$;

3. $p(A \cup \bar{A}) = p(A) + p(\bar{A}) = 1$, so that $p(A) \leq 1$ for any A in S ;

4. if $A \subset B$ then $p(B \setminus A) = p(B) - p(A)$;

Axiom (iii) is known as *countable additivity* and it is rejected by a school of probabilists who replace the countable additivity by finite additivity.

Consider now the two events A and B . If we computed $p(A \cup B)$ as $p(A) + p(B)$ we would obtain $p(A \cup B) = 1.1$ that exceeds 1. The error here is that, in computing $p(A \cup B)$ as $p(A) + p(B)$, the event A, B is counted twice. Indeed, we can decompose A into $(A \setminus B) \cup (A, B)$ and similarly B into $(A, B) \cup (B \setminus A)$. Since the intersection $(A \setminus B), (A, B) = \emptyset$, the events $(A \setminus B)$ and (A, B) are exclusive and there follows, from item 3 in Theorem 2.1, that $p(A) = p(A \setminus B) + p(A, B)$, and similarly $p(B) = p(A, B) + p(B \setminus A)$. The event $A \cup B$ is given by $(A \setminus B) \cup (A, B) \cup (B \setminus A)$, and the three events are exclusive. Thus, from property (2.7) we have $p(A \cup B) = p(A \setminus B) + p(A, B) + p(B \setminus A) = p(A \setminus B) + p(A, B) + p(B \setminus A) + p(A, B) - p(A, B) = p(A) + p(B) - p(A, B)$. The rule derived in this example holds in general:

$$p(A \cup B) = p(A) + p(B) - p(A, B)$$

SAMPLE AND EVENT SPACES

Sample Space -

A *probabilistic* (or *statistical*!) experiment has the following characteristics:

- (a) the set of all possible outcomes of the experiment can be described;
- (b) the outcome of the experiment cannot be predicted with certainty prior to the performance of the experiment.

The set of all possible outcomes (or sample points) of the experiment is called the *sample space* and is denoted by S . For a given experiment it may be possible to define several sample spaces.

Example For the experiment of tossing a coin three times, we could define

(a) $S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$ each outcome being an ordered sequence of results; or

(b) $S = \{0, 1, 2, 3\}$

each outcome being a possible value for the number of heads obtained.

If S consists of a list of outcomes (finite or infinite in number), S is a *discrete sample space*. Examples

i. Tossing a die: $S = \{1, 2, 3, 4, 5, 6\}$

ii. Tossing a coin until the first head appears: $S = \{H, TH, TTH, TTTH, \dots\}$

Otherwise S is an *uncountable sample space*. In particular, if S belongs to a Euclidean space (e.g. real line, plane), S is a *continuous sample space*.

Example Lifetime of an electronic device: $S = \{t : 0 \leq t < \infty\}$

Event Space -

Events - A specified collection of outcomes in S is called an *event*: i.e., any subset of S (including S itself) is an event. When the experiment is performed, an event A occurs if the outcome is a member of A .

Example In tossing a die once, let the event A be the occurrence of an even number: i.e., $A = \{2, 4, 6\}$. If a 2 or 4 or 6 is obtained when the die is tossed, event A occurs.

The event S is called the *certain event*, since some member of S must occur. A single outcome is called an *elementary event*. If an event contains no outcomes, it is called the *impossible* or *null event* and is denoted by \emptyset .

Combination of events - Since events are sets, they may be combined using the notation of set theory: Venn diagrams are useful for exhibiting definitions and results, and you should draw such a diagram for each operation and identity introduced below.

[In the following, A, B, C, A_1, \dots, A_n are events in the event space \mathcal{F} (discussed below), and are therefore subsets of the sample space S .

The *union* of A and B , denoted by $A \cup B$, is the event 'either A or B , or both'.

The *intersection* of A and B , denoted by $A \cap B$, is the event 'both A and B '.

The union and intersection operations are *commutative*, i.e.

$$A \cup B = B \cup A, \quad A \cap B = B \cap A, \quad (2.8)$$

associative, i.e.

$$\begin{aligned} A \cup (B \cup C) &= (A \cup B) \cup C = A \cup B \cup C, \\ A \cap (B \cap C) &= (A \cap B) \cap C = A \cap B \cap C, \end{aligned} \quad (2.9)$$

and *distributive*:

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned} \quad (2.10)$$

If A is a *subset* of B , denoted by $A \subset B$, then $A \cup B = B$ and $A \cap B = A$.

The *difference* of A and B , denoted by $A \setminus B$, is the event ' A but not B '.

The *complement* of A , denoted by \overline{A} , is the event 'not A '.

The complement operation has the properties:

$$A \cup \overline{A} = S, \quad A \cap \overline{A} = \emptyset, \quad \overline{(\overline{A})} = A, \quad (2.11)$$

$$\text{and } \overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}. \quad (2.12)$$

Note also that $A \setminus B = A \cap \overline{B}$. Hence use of the difference symbol can be avoided if desired (and will be in our discussion).

Two events A and B are termed *mutually exclusive* if $A \cap B = \emptyset$.

Two events A and B are termed *exhaustive* if $A \cup B = S$.

The above results may be generalized to combinations of n events: thus $A_1 \cup A_2 \cup \dots \cup A_n$ is the event 'at least one of A_1, A_2, \dots, A_n '.

$A_1 \cap A_2 \cap \dots \cap A_n$ is the event 'all of A_1, A_2, \dots, A_n '.

$$\begin{aligned} &A_1 \cap (A_2 \cup A_3 \cup \dots \cup A_n) \\ &= (A_1 \cap A_2) \cup (A_1 \cap A_3) \cup \dots \cup (A_1 \cap A_n) \end{aligned} \quad (2.13)$$

$$\begin{aligned} &A_1 \cup (A_2 \cap A_3 \cap \dots \cap A_n) \\ &= (A_1 \cup A_2) \cap (A_1 \cup A_3) \cap \dots \cap (A_1 \cup A_n) \end{aligned} \quad (2.14)$$

$$\begin{aligned} \overline{A_1 \cup A_2 \cup \dots \cup A_n} &= \overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_n} \quad \text{or} \\ \overline{\bigcup_{i=1}^n A_i} &= \bigcap_{i=1}^n \overline{A_i} \end{aligned} \quad (2.15)$$

$$\begin{aligned} \overline{A_1 \cap A_2 \cap \dots \cap A_n} &= \overline{A_1} \cup \overline{A_2} \cup \dots \cup \overline{A_n} \quad \text{or} \\ \overline{\bigcap_{i=1}^n A_i} &= \bigcup_{i=1}^n \overline{A_i} \end{aligned} \quad (2.16)$$

(The last two results are known as de Morgan's Laws - see e.g. Ross for proofs.)

The events A_1, A_2, \dots, A_n are termed *mutually exclusive* if $A_i \cap A_j = \emptyset$ for all $i \neq j$.

The events A_1, A_2, \dots, A_n are termed *exhaustive* if $A_1 \cup A_2 \cup \dots \cup A_n = S$.

If the events A_1, A_2, \dots, A_n are both mutually exclusive and exhaustive, they are called a *partition* of S .

Event space - The collection of all subsets of S may be too large for probabilities to be assigned reasonably to all its members. This suggests the concept of *event space*.

A collection \mathcal{F} of subsets of the sample space is called an *event space* (or σ -field) if (a) the certain event S and the impossible event \emptyset belong to \mathcal{F} ;

(b) if $A \in \mathcal{F}$, then $\bar{A} \in \mathcal{F}$;

(c) if $A_1, A_2, \dots \in \mathcal{F}$, then $A_1 \cup A_2 \cup \dots \in \mathcal{F}$,

i.e. \mathcal{F} is closed under the operation of taking countable unions. It is readily shown that, if $A_1, A_2, \dots \in \mathcal{F}$, then $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$. For $\bigcap_{i=1}^{\infty} \bar{A}_i = \overline{\bigcup_{i=1}^{\infty} A_i} \in \mathcal{F}$ (invoking properties (b) and (c) of \mathcal{F} , and the result follows from property (b)).

For a *finite* sample space, we normally use the collection of all subsets of S (the *power set* of S) as the event space. For $S = \{-\infty, \infty\}$ (or a subset of the real line), the collection of sets containing all one-point sets and all well-defined intervals is an event space.

CONDITIONAL PROBABILITY

The probabilities considered so far are unconditional probabilities. In some situations, however, we may be interested in the probability of an event given the occurrence of some other event. For instance, the probability of R : "Tomorrow, January 16th, it will rain in Amherst" would change, if we happened to know that tomorrow is a cloudy day. Formally, if we denote by C the event "Tomorrow, 16th of January, will be cloudy", assuming the occurrence of C is equivalent to restricting our sample space, because other events as S (sunny day) are ruled out. We thus need to recompute the probability of R by taking into account this new piece of information. This is formally done by considering the *conditional probability* of R given that C occurs. This event is denoted by $R|C$.

Consider the events A and B . If we limit the scenario of possible events to A , the occurrence of B would be restricted to A . If we knew that A occurs, we would then deduce $p(B|A) = p(A, B)$. However, since $p(A) < 1$, we can only state that $p(B|A) = kp(A, B)$, where k is proportionality constant that accounts for the uncertainty in the occurrence of A . Clearly, we have $p(A|A) = 1$ and also $p(A|A) = kp(A, A) = kp(A)$. From this, we deduce that $k = 1/p(A)$ and the conditional probability is thus defined as follows.

Definition 2.1 (Conditional Probability) Let A and B events in S , and suppose that $p(A) > 0$. The conditional probability of B given A is:

$$p(B|A) = \frac{p(B, A)}{p(A)}. \tag{2.17}$$

To emphasize that $p(A)$ is unconditional, $p(A)$ is called *marginal probability*.

Example 2.1 (Conditional Probability) Consider choosing a card from a well-shuffled standard deck of 52 playing cards. The probability that the first card extracted is an ace is clearly $4/52$. Suppose that, after the first extraction, the card is not reinserted in the deck. What is the probability that the second card is an ace, given that the first card is an ace? Let A be the event that the first card is an ace, and let B be the event that the second card is an ace. The probability of A, B is

$$\frac{12}{2652} = \frac{\text{possible pairs of aces}}{\text{all possible pairs of cards}}$$

and $p(A) = 4/52$. On using (2.17) we have

$$p(B|A) = \frac{\frac{12}{2652}}{\frac{4}{52}} = \frac{3}{51}.$$

Indeed, there are three aces left in a deck of 51 cards.

From the definition of conditional probability in (2.17), we derive the probability of the intersection of two events, called their *joint probability* in terms of conditional and marginal probabilities:

$$p(A, B) = p(B|A)p(A). \tag{2.18}$$

This rule can be applied to a larger number of events and produces the *multiplication rule* or *factorization rule*.

Definition 2.2 (Multiplication Rule) The joint probability of a set of events A_1, A_2, \dots, A_n can be expressed as $p(A_1, A_2, \dots, A_n) = p(A_1)p(A_2|A_1)p(A_3|A_1, A_2) \dots p(A_n|A_1, A_2, \dots, A_{n-1})$

Consider again the events A and B . The events A and \bar{A} form a partition of S , so that we can decompose B into the union of the two exclusive events A, B and \bar{A}, B . Thus, if we use (4) and the third axiom of we have:

$$p(B) = p(A, B) + p(\bar{A}, B) = p(B|A)p(A) + p(B|\bar{A})p(\bar{A}). \tag{2.19}$$

Formula (2.19) is known as the *Total Probability Theorem* and expresses the marginal probability of B as a weighted average of the conditional probabilities $p(B|A)$ and $p(B|\bar{A})$ with weights given by $p(A)$ and $p(\bar{A})$. The importance of the Total Probability Theorem is that, sometimes, expressing conditional probabilities can be easier than expressing marginal probabilities, and (2.19) can be used to “break down” an event in more specific events, on which a more precise knowledge is available. Suppose, as an example, that B is the event that the result of a test to diagnose the presence of a disease A is positive. Quantifying the incidence of false positive ($B|\bar{A}$) and false negative ($\bar{B}|A$) can be easier than quantifying the marginal probability of B . If, further, the incidence rate of A is known, then (2.19) can be used to derive $p(B)$.

The multiplication rule and the Total probability theorem can be extended to conditional probabilities. So,

$$p(A_1, A_2, \dots, A_n|C) = p(A_1|C)p(A_2|A_1, C)p(A_3|A_1, A_2, C) \\ \dots p(A_n|A_1, A_2, \dots, A_n, C)$$

and

$$p(B|C) = p(B|A, C)p(A|C) + p(B|\bar{A}, C)p(\bar{A}|C).$$

REFERENCES

- Carlos Al'os-Ferrer and Ana B. Ania (2005). The evolutionary stability of perfectly competitive behavior. *Economic Theory*, 26: pp. 497–516.
- Casella, G., and Berger, R. L. (2000). *Statistical Inference*. Duxbury Press, Belmont, Ca.
- Cover, T. M., and Thomas, M. (2001). *Elements of Information Theory*. Wiley, New York, NY.
- Evans, M., and Swartz, T. (2005). Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. *Statistical Science*, 10, pp. 254–272.
- Karr, A. F. (2002). *Probability*. Springer-Verlag, New York.
- Whittaker, J. (2000). *Graphical Models in Applied Multivariate Statistics*. Wiley, New York, NY.

Corresponding Author

Rajkumar*

Author Designation

E-Mail – ahujarajkumar22@gmail.com