Contents

Chapter 1

Probability

Note: These lecture notes aim to present a clear and crisp presentation of some topics in Probability and Statistics. Comments/suggestions are welcome on the e-mail: sukuyd@gmail.com to Dr. Suresh Kumar.

1.1 Introduction

Engineers and scientists are constantly exposed to collections of facts, or data. The discipline of statistics provides methods for organizing/summarizing data and for drawing conclusions based on information contained in the data.

A statistical investigation is typically focused on a well-defined collection of objects constituting a **population** of interest. When desired information is available for all objects in the population, we have what is called a **census**. But obtaining census of large size populations is usually impractical or infeasible due to time, money, resources and many other constraints. In such situations, a subset of the population - **a sample** - is selected in some prescribed manner. e.g., Suppose we wish to calculate average income per person in India. So we need to collect the information about the income of each person in India which is impractical. So sampling becomes essential in such situations.

We are usually interested only in certain characteristics of the objects in a population. A characteristic may be **categorical** or it may be **numerical** in nature. For example, we may be interested to collect the information about gender and age of engineering graduates. Here, the value of the characteristic 'gender' is a category, whereas the value of the characteristic 'age' is a number.

A **variable** is any characteristic whose value may change from one object to another in the population. Data arise from observations either on a single variable or simultaneously on two or more variables. A univariate data set consists of observations on a single variable. For example, we might determine the type of transmission, automatic (A) or manual (M), on each of ten automobiles recently purchased at a certain dealership, resulting in the categorical univariate data set: M A A A M A A M A A

The following sample of lifetimes (hours) of brand D batteries put to a certain use is a numerical univariate data set: 5.6 5.1 6.2 6.0 5.8 6.5 5.8 5.5

We have **bivariate** data when observations are made on each of two variables. Our data set might consist of a (height, weight) pair for each basketball player on a team, with the first observation as (72, 168), the second as (75, 212), and so on. In general, **Multivariate** data arise when observations are made on more than one variable (so bivariate is a special case of multivariate).

Descriptive Statistics belongs to the data analysis in the cases where the data set size is manageable and can be analysed analytically or graphically. Some of these methods are graphical in nature; the construction of histograms, boxplots, and scatter plots are primary examples. Other descriptive methods involve calculation of numerical summary measures, such as means, standard deviations, and correlation coefficients.

On the other hand, **Inferential Statistics** is applied where the entire data set (population) can not be analysed at one go or as a whole. So we draw a sample (a small or manageable portion) from the population. Then we analyse the sample for the characteristic of interest and try to infer the same about the population. For example, when you cook rice, you take out few grains and crush them to see whether the rice is properly cooked. Similarly, survey polls prior to voting in elections, TRP ratings of TV channel shows etc are samples based and therefore belong to the inferential statistics.

The discipline of **probability** forms a bridge between the descriptive and inferential techniques. Mastery of probability leads to a better understanding of how inferential procedures are developed and used, how statistical conclusions can be translated into everyday language and interpreted, and when and where pitfalls can occur in applying the methods. Probability and statistics both deal with questions involving populations and samples, but do so in an “inverse manner” to one another. The relationship between the two disciplines can be summarized by saying that probability reasons from the population to the sample (deductive reasoning), whereas inferential statistics reasons from the sample to the population (inductive reasoning).

Before we can understand what a particular sample can tell us about the population, we should first understand the uncertainty associated with taking a sample from a given population. This is why we study probability before statistics. So statistics is fundamentally based on the theory of probability. So first we discuss the theory of probability, and then we shall move to the statistical methods.

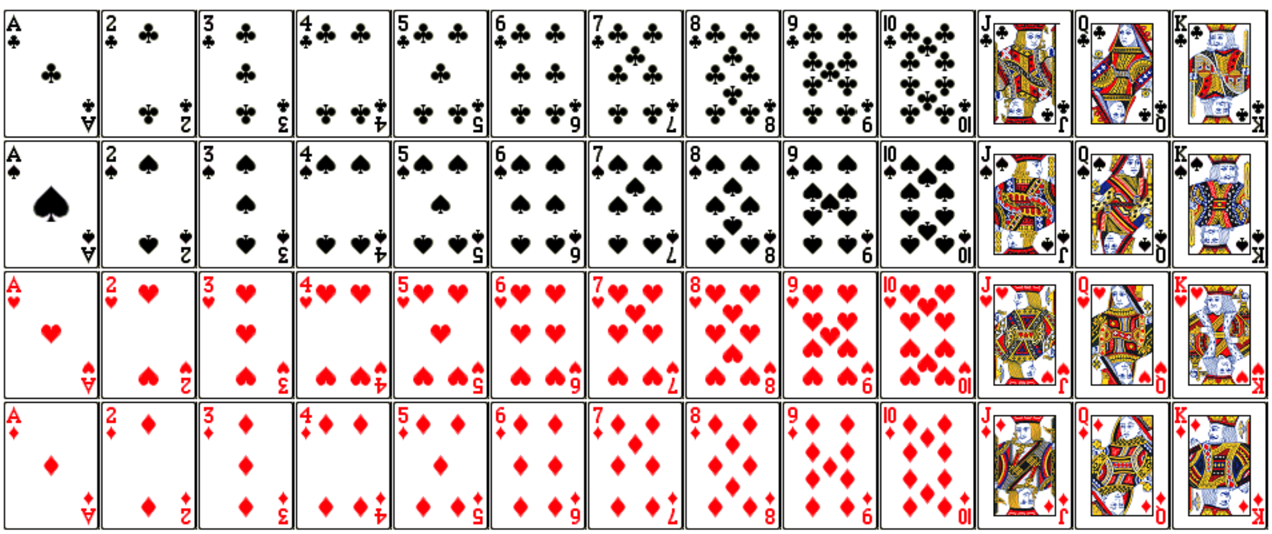

Figure 1.1: 52 cards of a deck: The four rows are the four suits. The clubs are all in the first row, followed by the spades, then the hearts, and last the diamonds. Among the 13 kinds, we find the numbers 2 through 10, and four other kinds. The A stands for ace, the J for jack, the Q for queen, and the K for king. The jack, queen, and king are often referred to as face cards.

1.2 Theory of Probability

1.2.1 Definitions

Random Experiment

An experiment whose outcome or result is random, that is, is not known before the experiment, is called random experiment. eg.

(i) Tossing a fair coin
(ii) Rolling a fair die
(iii) Drawing a card from a well-shuffled pack of cards

all are random experiments.

Sample Space

Set of all possible outcomes is called sample space of the random experiment and is usually denoted by $\mathcal{S}$.

Ex. When a fair coin is tossed or flipped, it either shows the head H or the tail T. So sample space of this experiment is

$$\mathcal{S} = \{H, T\}.$$

Ex. Consider the experiment of rolling a die. If we are interested in the number that shows on the top face, the sample space is

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}.$$

If we are interested only in whether the number is even or odd, the sample space is simply

$$\mathcal{S} = \{\text{even}, \text{odd}\}.$$

In general, it is desirable to use the sample space that gives the most information concerning the outcomes of the experiment. In some experiments, it is helpful to list the elements of the sample space systematically by means of a tree diagram.

Ex. An experiment consists of flipping a coin and then flipping it a second time if a head occurs. If a tail occurs on the first flip, then a die is rolled once. To list the elements of the sample space providing the most information, we construct the tree diagram in Figure 1.

By proceeding along all paths, we see that the sample space is

$$\mathcal{S} = \{HH, HT, T1, T2, T3, T4, T5, T6\}.$$

Ex. Suppose that three items are selected at random from a manufacturing process. Each item is inspected and classified defective, D, or nondefective, N. To list the elements of the sample space providing the most information, we construct the tree diagram of Figure 2.

By proceeding along all paths, we see that the sample space is

$$\mathcal{S} = \{DDD, DDN, DND, DNN, NDD, NDN, NND, NNN\}.$$

Sample spaces with a large or infinite number of sample points are best described by a statement or rule method. For example, if $\mathcal{S}$ is the set of all points (x, y) on the boundary or the interior of a circle of radius 2 with center at the origin, we write the rule

$$\mathcal{S} = \{(x, y) | x^2 + y^2 \leq 4\}.$$

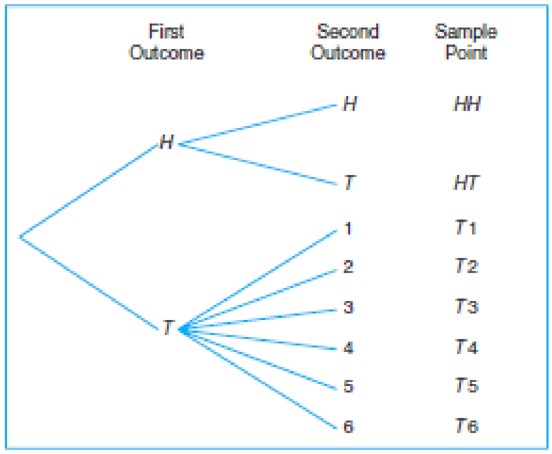

Figure 1.2: Tree diagram

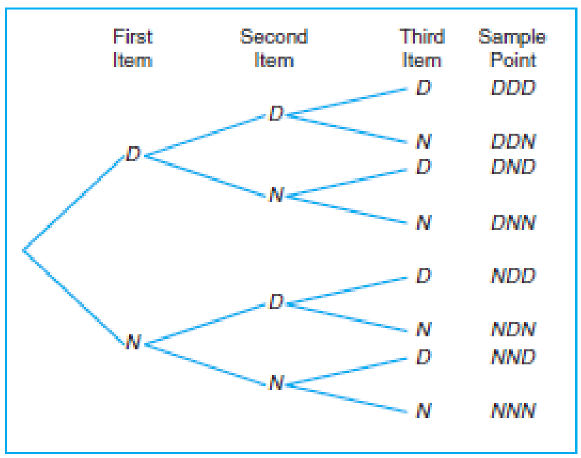

Figure 1.3: Tree diagram

Consider a random experiment which consists of a series of tosses of a fair coin till the head appears. The head may turn up in the first toss or the second toss and so forth. Thus, the sample space is infinite and can be written as

$$\mathcal{S} = \{H, TH, TTH, TTTH, \text{ ....}\}.$$

Events

Any subset of sample space is called an event. eg. If $S = \{H, T\}$, then the sets ϕ, $\{H\}$, $\{T\}$ and $\{H, T\}$ all are events. The event ϕ is called impossible event as it does not happen. The event $\{H, T\}$ is called sure event as we certainly get either head or tail in the toss of a fair coin.

Elementary and Compound Events

Singleton subsets of sample space $\mathcal{S}$ are called elementary events. The subsets of $\mathcal{S}$ containing more than one element are known as compound events. eg. The singleton sets $\{H\}$ and $\{T\}$ are called elementary events while $\{H, T\}$ is a compound event.

Equally Likely Events

The elementary events of a sample space are said to be equally likely if each one of them has same chance of occurring. eg. The elementary events $\{H\}$ and $\{T\}$ in the sample space of the toss of a fair coin are equally likely because both have same chance of occurring.

Mutually Exclusive and Exhaustive Events

Two events are said to be mutually exclusive If happening of one event precludes the happening of the other. eg. The events $\{H\}$ and $\{T\}$ in the sample space of the toss of a fair coin are mutually exclusive because both can not occur together. Similarly, more than two events say A_1, A_2,....,A_n are mutually exclusive if any two of these can not occur together, that is, $A_i \cap A_j = \phi$ for $i \neq j$, where $i, j \in \{1, 2, ..., n\}$. Further, the mutually exclusive events in a sample space are exhaustive if their union is equal to the sample space. eg. The events $\{H\}$ and $\{T\}$ in the sample space of the toss of a fair coin are mutually exclusive and exhaustive.

Combination of Events

If A and B are any two events in a sample space S, then the event $A \cup B$ implies either A or B or both; $A \cap B$ implies both A and B; $A - B$ implies A but not B; $\bar{A}$ implies not A, that is, $\bar{A} = S - A$.

eg. Let S be sample space in a roll of a fair die. Then $S = \{1, 2, 3, 4, 5, 6\}$. Let A be the event of getting an even number and B be the event of getting a number greater than 3. Then $A = \{2, 4, 6\}$ and $B = \{4, 5, 6\}$. So $A \cup B = \{2, 4, 5, 6\}$, $A \cap B = \{4\}$, $A - B = \{2\}$ and $\bar{A} = \{1, 3, 5\}$.

Counting Sample Points

One of the problems that the statistician must consider and attempt to evaluate is the element of chance associated with the occurrence of certain events when an experiment is performed. These problems belong in the field of probability. In many cases, we shall be able to solve a probability problem by counting the number of points in the sample space without actually listing each element. In this regard, the following rules of counting are useful.

If an operation can be performed in n_1 ways, and if for each of these ways a second operation can be performed in n_2 ways, then the two operations can be performed together in $n_1 n_2$ ways.

This is the **fundamental principle of counting**, often referred to as the multiplication rule.

Ex. How many sample points are there in the sample space when a pair of dice is thrown once?
Sol. The first die can land face-up in any one of $n_1 = 6$ ways. For each of these 6 ways, the second die can also land face-up in $n_2 = 6$ ways. Therefore, the pair of dice can land in $n_1 n_2 = (6)(6) = 36$ possible ways. So the sample space carries 36 points given by

$$S = \{(1, 1), (1, 2), ..., (1, 6), (2, 1), (2, 2), ...(2, 6), ..., (6, 6)\}.$$

Ex. A developer of a new subdivision offers prospective home buyers a choice of Tudor, rustic, colonial, and traditional exterior styling in ranch, two-story, and split-level floor plans. In how many different ways can a buyer order one of these homes?
Sol. Since $n_1 = 4$ and $n_2 = 3$, a buyer must choose from $n_1 n_2 = (4)(3) = 12$ possible homes. It can also be observed/verified from the tree diagram in Figure 3.
Ex. If a 22-member club needs to elect a chair and a treasurer, how many different ways can these two to be elected?

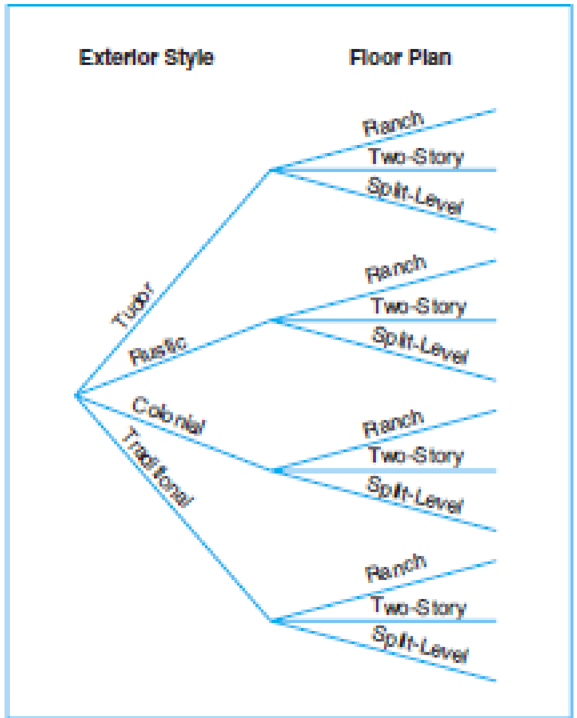

Figure 1.4: Tree diagram

Sol. For the chair position, there are 22 total possibilities. For each of those 22 possibilities, there are 21 possibilities to elect the treasurer. Using the multiplication rule, we obtain $n_1 \times n_2 = 22 \times 21 = 462$ different ways.

Note that the fundamental principle of counting or the multiplication rule may be extended to cover any number of operations. Suppose, for instance, that a customer wishes to buy a new cell phone and can choose from $n_1 = 5$ brands, $n_2 = 5$ sets of capability, and $n_3 = 4$ colors. These three classifications result in $n_1 n_2 n_3 = (5)(5)(4) = 100$ different ways for a customer to order one of these phones.

Ex. A person is going to assemble a computer by himself. He has the choice of chips from two brands, a hard drive from four, memory from three, and an accessory bundle from five local stores. How many different ways can the person order the parts?
Sol. Since $n_1 = 2, n_2 = 4, n_3 = 3$, and $n_4 = 5$, there are

$$n_1 \times n_2 \times n_3 \times n_4 = 2 \times 4 \times 3 \times 5 = 120$$

different ways to order the parts.

Ex. How many even four-digit numbers can be formed from the digits 0, 1, 2, 5, 6, and 9 if each digit can be used only once?
Sol. Answer is 156. (Please try yourself)

Permutations of distinct objects

Frequently, we are interested in a sample space that contains as elements all possible orders or arrangements of a group of objects. For example, we may want to know how many different arrangements are possible for sitting 6 people around a table, or we may ask how many different orders are possible for drawing 2 lottery tickets from a total of 20. The different arrangements are called permutations.

Ex. Consider the three letters a, b, and c. The possible permutations are abc, acb, bac, bca, cab, and cba. Thus, we see that there are 6 distinct arrangements. Using the multiplication rule, we could arrive at the answer 6 without actually listing the different orders by the following arguments: There are $n_1 = 3$ choices for the first position. No matter which letter is chosen, there are always $n_2 = 2$ choices for the second position. No matter which two letters are chosen for the first two positions, there is only $n_3 = 1$ choice for the last position, giving a total of $n_1 n_2 n_3 = (3)(2)(1) = 6$ permutations. In general, n distinct objects can be arranged in

$$n(n-1)(n-2)......(3)(2)(1) \text{ ways.}$$

For any non-negative integer n, the expression $n(n-1)(n-2)......(3)(2)(1)$ is denoted by $n!$, called "n factorial" with the special case definition $0! = 1$. Thus, the number of permutations of n distinct objects is $n!$.

The number of permutations of the four letters a, b, c, and d will be $4! = 24$. Now consider the number of permutations that are possible by taking two letters at a time from four. These would be ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, and dc. Using multiplication rule again, we have two positions to fill, with $n_1 = 4$ choices for the first and then $n_2 = 3$ choices for the second, for a total of $n_1 n_2 = (4)(3) = 12$ permutations. In general, n distinct objects taken r at a time can be arranged in $n(n-1)(n-2)(n-r+1)$ ways. We will represent this product by the symbol $P_{r,n}$ (also denoted by the symbol ${}^nP_r$), and read as "Permutations of n objects taken r at a time". So we have

$$P_{r,n} = n(n-1)(n-2)(n-r+1)$$

$$= \frac{n(n-1)(n-2)(n-r+1) \times (n-r)!}{(n-r)!} = \frac{n!}{(n-r)!}.$$

Thus the number of permutations of n distinct objects taken r at a time is

$$P_{r,n} = \frac{n!}{(n-r)!}.$$

Ex. In one year, three awards (research, teaching, and service) will be given to a class of 25 graduate students in a statistics department. If each student can receive at most one award, how many possible selections are there?
Sol. $P_{3,25} = \frac{25!}{22!} = (25)(24)(23) = 13800$.

Ex. A president and a treasurer are to be chosen from a student club consisting of 50 people. How many different choices of officers are possible if (a) there are no restrictions; (b) A will serve only if he is president; (c) B and C will serve together or not at all; (d) D and E will not serve together?
Sol. (a) $P_{2,50} = 2450$ (b) $49 + P_{2,49} = 2352$ (c) $2 + P_{2,48} = 2258$ (d) $2450 - 2 = 2440$.

Permutations of repeated objects

So far we have considered permutations of distinct objects. That is, all the objects were completely different or distinguishable. Obviously, if the letters b and c are both equal to x, then the 6 permutations of the letters a, b, and c become axx, axx, xax, xax, xxa, and xxa, of which only 3 are distinct. Therefore, with 3 letters, 2 being the same, we have $3!/2! = 3$ distinct permutations.

Likewise, With 4 different letters a, b, c, and d, we have 24 distinct permutations. If we let a = b = x and c = d = y, we can list only the following distinct permutations: xxyy, xyxy, yxxy, yyxx, xyyx, and yxyx. Thus, we have $4!/(2!2!) = 6$ distinct permutations.

In general, the number of distinct permutations of n things of which n_1 are of one kind, n_2 of a second kind, . . . , n_k of a kth kind is

$$\frac{n!}{n_1!\, n_2! \,.....\, n_k!}.$$

Ex. How many different letter arrangements can be made from the letters in the word STATISTICS ?
Sol. $\dfrac{10!}{3!\,3!\,2!} = 50400.$

Ex. In how many ways can 7 graduate students be assigned to 1 triple and 2 double hotel rooms during a conference?
Sol. $\dfrac{7!}{3!\,2!\,2!} = 210.$

Combinations

In many problems, we are interested in the number of ways of selecting r objects from n objects without regard to order or the r objects are of same kind in the sense of permutations. These selections are called combinations. The number of combinations from n objects taking r objects at a time is denoted by ${}^nC_r$ or $C_{r,n}$ or $\binom{n}{r}$, and is given by

$$C_{r,n} = \binom{n}{r} = \frac{P_{r,n}}{r!} = \frac{n!}{(n-r)!\,r!}.$$

Ex. In how many ways can 2 players be selected from a group of 5 players?

Sol. $\dfrac{5!}{(5-2)!\,2!} = 10.$

1.2.2 Axioms of Probability

(i) The probability of any event A, denoted by $P(A)$, is to be assigned a number in the closed interval $[0,1]$, that is, $0 \leq P(A) \leq 1$.
(ii) $P(\phi) = 0$ and $P(S) = 1$.
(iii) If A and B are mutually exclusive or disjoint events, then $P(A \cup B) = P(A) + P(B)$.

These are known as axioms[1] of the theory of probability. The axiom (iii) may be assumed for arbitrary union of disjoint sets.

Deductions:

One may easily deduce the following results from the above axioms:

(i) If A and B are any two events, then
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$. This is called **law of addition of probabilities**.

Proof: We write $A = (A - B) \cup (A \cap B)$, the union of two disjoint sets. So by axiom (iii), we have $P(A) = P(A - B) + P(A \cap B)$ or $P(A - B) = P(A) - P(A \cap B)$. Also, we can write $A \cup B = (A - B) \cup B$, the union of two disjoint sets. So we have

[1] Axioms are mathematical statements or assumptions without proof, which form the basis of the logical development of a theory.

$P(A \cup B) = P(A - B) + P(B) = P(A) - P(A \cap B) + P(B)$.

Likewise, for three events A, B and C, we may deduce the addition rule given by
$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C)$.

(ii) $P(A') = 1 - P(A)$. It follows from the fact that A and A' are disjoint and $A \cup A' = \mathcal{S}$ with $P(\mathcal{S}) = 1$.

(iii) If A is subset of B, then $P(A) \leq P(B)$. For, A being subset of B, the sets $B - A$ and A are disjoint with $B = (B - A) \cup A$. It follows that
$P(B) - P(A) = P((B - A) \cup A) - P(A) = P(B - A) + P(A) - P(A) = P(B - A) \geq 0$.

Ex. In a certain residential suburb, 60% of all households get Internet service from the local cable company, 80% get television service from that company, and 50% get both services from that company. If a household is randomly selected, what is the probability that it gets at least one of these two services from the company, and what is the probability that it gets exactly one of these services from the company?
Sol. Let A and B be the events that a household gets internet and television services, respectively, from the local company. So $P(A) = 0.6$, $P(B) = 0.8$ and $P(A \cap B) = 0.5$. Then the probability that a household gets at least one of these two services from the company is given by

$$P(A \cup B) = 0.6 + 0.8 - 0.5 = 0.9$$

The probability of getting exactly one of the two services is

$$P(A \cap B') + P(A' \cap B) = 0.1 + 0.3 = 0.4$$

1.2.3 Assigning Probabilities to Events

In the above example, the probabilities are already assigned to the events of interest. If the probabilities are not assigned to the events of interest, then we need to do it logically. There exist different approaches to do so. For instance, in relative frequency approach, we assign the probability to an event on the basis of its frequency in the repeated random experiment. For example, if we toss a fair coin 10 times, and head appears 7 times, then the probability of getting head in the 11th toss is 7/10. On the other hand, in classical approach, we assign equal weight or probability to the equally likely outcomes. For example, the outcomes head (H) and tail (T) in the toss of a fair coin are equally likely. So we assign equal probability, say w to both the events, that is, $P(\{H\}) = w$ and $P(\{T\}) = w$. Also in view of axioms (ii) and (iii), $\mathcal{S} = \{H, T\} = \{H\} \cup \{T\}$ implies $1 = P(\mathcal{S}) = P(\{H\}) + P(\{T\}) = w + w = 2w$. It follows that $P(\{H\}) = 1/2$ and $P(\{T\}) = 1/2$. Thus, in the classical approach, probability of getting head is 1/2.

Ex. A die is loaded in such a way that an even number is twice as likely to occur as an odd number. If E is the event that a number less than 4 occurs on a single toss of the die, find $P(E)$.
Sol. The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. We assign a probability of w to each odd number and a probability of $2w$ to each even number. Since the sum of the probabilities must be 1, we have $9w = 1$ or $w = 1/9$. Hence, probabilities of 1/9 and 2/9 are assigned to each odd and even number, respectively. Therefore, $E = \{1, 2, 3\}$ and $P(E) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \dfrac{4}{9}$.

1.2.4 Classical Formula of Probability

Let $\mathcal{S} = \{a_1, a_2, ..., a_m\}$ be sample space of a random experiment, where the m outcomes $a_1, a_2, ..., a_m$ are equally likely. In the classical approach, we assign equal probability $1/m$ to each outcome or elementary event. If A is any event with k elements in $\mathcal{S}$, then A is the union of k elementary events. So probability of A is given by

$$P(A) = \frac{k}{m} = \frac{\text{Numer of elements in } A}{\text{Number of elements in } \mathcal{S}} = \frac{n(A)}{n(\mathcal{S})}.$$

Ex. Find the probability of getting exactly two heads in toss of two fair coins?

Sol. If $\mathcal{S}$ is sample space for toss of two fair coins, then $\mathcal{S} = \{HH, HT, TH, TT\}$. The coins being fair, here all the four outcomes are equally likely. Let A be the event of getting two heads. Then $A = \{HH\}$, and therefore $P(A) = 1/4$.

Ex. In a poker hand consisting of 5 cards, find the probability of holding 2 aces and 3 jacks.

Sol. $\dfrac{\binom{4}{2}\binom{4}{3}}{\binom{52}{5}} = 0.00009.$

Note that the classical approach is applicable in the cases (such as the above example) where it is reasonable to assume that all possible outcomes are equally likely. The probability assigned to an event through classical approach is the accurate probability.

Ex. From a pack of well shuffled cards, one card is drawn. Find the probability that the card is either a king or an ace.

Sol. $4/52 + 4/52 = 2/13$

Ex. Two dice are rolled once. Find the probability of getting an even number on the first die or a total of 8.

Sol. $18/36 + 5/36 - 3/36 = 5/9$

Ex. John is going to graduate from an industrial engineering department in a university by the end of the semester. After being interviewed at two companies he likes, he assesses that his probability of getting an offer from company A is 0.8, and his probability of getting an offer from company B is 0.6. If he believes that the probability that he will get offers from both companies is 0.5, what is the probability that he will get at least one offer from these two companies?

Sol. $0.8 + 0.6 - 0.5 = 0.9$

Ex. If the probabilities are, respectively, 0.09, 0.15, 0.21, and 0.23 that a person purchasing a new automobile will choose the color green, white, red, or blue, what is the probability that a given buyer will purchase a new automobile that comes in one of those colors?

Sol. $0.09 + 0.15 + 0.21 + 0.23 = 0.68$

Ex. If the probabilities that an automobile mechanic will service 3, 4, 5, 6, 7, or 8 or more cars on any given workday are, respectively, 0.12, 0.19, 0.28, 0.24, 0.10, and 0.07, what is the probability that he will service at least 5 cars on his next day at work?

Sol. $1 - (0.12 + 0.19) = 0.69$

Ex. Assuming 365 days in year, find the probability that among 30 students in a classroom, at least two will have the same birthday.

Sol. $1 - \frac{P_{30,365}}{365^{30}}$.

Ex. In a game, each player chooses six numbers between one and forty nine. If these numbers all match the six winning numbers, then the player wins the first prize. For the second prize, a player needs to match five numbers with the winning numbers. What are the probabilities of winning the first and second prizes?

Sol. (i) $\dfrac{C_{6,6}}{C_{6,49}}$ (ii) $\dfrac{C_{1,43}C_{5,6}}{C_{6,49}}$

1.2.5 Conditional Probability

Suppose a bag contains 10 Blue, 15 Yellow and 20 Green balls where all balls are identical except for the color. Let B be the event of drawing a Blue ball from the bag. Then $P(B) = 10/45 = 2/9$. Now suppose we are told after the ball has been drawn that the ball drawn is not Green. Let this event be A. So the event A carries 10 Blue and 15 Yellow balls, and the event $B \cap A$ carries 10 Blue balls. It implies that $P(A) = 25/45$ and $P(B \cap A) = 10/45$. Now, considering the extra information/condition or happening of the event A, we need to calculate the revised/conditional probability of B, that we denote by $P(B|A)$. Since the ball is not Green, the revised sample space carries 25 balls (10 Blue and 15 Yellow). Hence, $P(B|A) = 10/25 = 2/5$. It can also be written as $P(B|A) = \frac{10}{25} = \frac{10/45}{25/45} = \dfrac{P(B \cap A)}{P(A)}$.

Obviously, more the overlapping of B with A, more is the value of $P(B|A)$. Therefore, $P(B|A) \propto P(B \cap A)$ or $P(B|A) = kP(B \cap A)$, k being some proportionality constant. Since $P(A|A) = 1$, so $1 = kP(A)$ and hence $P(B|A) = \dfrac{P(B \cap A)}{P(A)}$. Thus, formally the conditional probability is defined as follows.

For any two events A and B with $P(A) > 0$, the conditional probability of B given that A has occurred is defined by

$$P(B|A) = \frac{P(B \cap A)}{P(A)}.$$

Ex. The probability that a regularly scheduled flight departs on time is $P(D) = 0.83$; the probability that it arrives on time is $P(A) = 0.82$; and the probability that it departs and arrives on time is $P(D \cap A) = 0.78$. Find the probability that a plane (a) arrives on time, given that it departed on time, and (b) departed on time, given that it has arrived on time.

Sol. (a) $P(A/D) = P(D \cap A)/P(D) = 0.78/0.83 = 0.94$
(b) $P(D|A) = P(D \cap A)/P(A) = 0.78/0.82 = 0.95$

Ex. The concept of conditional probability has countless uses in both industrial and biomedical applications. Consider an industrial process in the textile industry in which strips of a particular type of cloth are being produced. These strips can be defective in two ways, length and nature of texture. For the case of the latter, the process of identification is very complicated. It is known from historical information on the process that 10% of strips fail the length test, 5% fail the texture test, and only 0.8% fail both tests. If a strip is selected randomly from the process and a quick measurement identifies it as failing the length test, what is the probability that it is texture defective?

Sol. (a) $P(T|L) = P(T \cap L)/P(L) = 0.008/0.1 = 0.08$. Thus, knowing the conditional probability provides considerably more information than merely knowing $P(T)$.

Ex. A die is rolled twice and the sum of the numbers appearing is noted to be 8. What is the probability that the number 5 has appeared at least once?

Sol. $P(A) = 11/36$, $P(B) = 5/36$, $P(A \cap B) = 2/36$, $P(A|B) = 2/5$.

1.2.6 Multiplication Rule

Suppose two events A and B happen in succession. Then the probability of happening of A and B is given by

$$P(A \cap B) = P(A)P(B|A).$$

This is known as multiplication rule of probabilities, and is simply the consequence of conditional probability rule. This rule can be extended to the case of three or more events happening in succession. For example, the probability of happening of three events A, B and C in succession is given by

$$P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B).$$

Here $P(A)$ is the probability of happening of the first event A; $P(B|A)$ is the probability of happening of the second event B given that the first event A has occurred, and $P(C|A \cap B)$ is the probability of happening of the third event C given that the first event A and second event B have occurred.

Ex. Suppose that we have a fuse box containing 20 fuses, of which 5 are defective. If 2 fuses are selected at random and removed from the box in succession without replacing the first, what is the probability that both fuses are defective?

Sol. We shall let A be the event that the first fuse is defective and B the event that the second fuse is defective; then we interpret $A \cap B$ as the event that A occurs and then B occurs after A has occurred. The probability of first removing a defective fuse is $P(A) = 1/4$; then the probability of removing a second defective fuse from the remaining 4 is $P(B|A) = 4/19$. Hence, $P(A \cap B) = P(A)P(B|A) = (1/4)(4/19) = 1/19$.

Ex. Two cards are drawn one after the other from a pack of well-shuffled 52 cards. Find the probability that both are spade cards if the first card is not replaced.

Sol. Let A be the event that the first drawn card is spade, and B be the event that the second drawn card is spade. So $P(A) = 13/52$, and $P(B|A) = 12/51$. So required probability is $P(A \cap B) = P(A)P(B|A) = (13/52)(12/51) = 1/17$.

Note. The event "Drawing two cards one by one without replacing the first one" is same as the event "Drawing two cards at one go". So in the above example, the probability of the event "both are spade cards" can also be calculated using combinations:

$$P(A \cap B) = \frac{\binom{13}{2}}{\binom{52}{2}} = \frac{13 \times 12}{52 \times 51} = \frac{1}{17}.$$

Ex. One bag contains 4 white balls and 3 black balls, and a second bag contains 3 white balls and 5 black balls. One ball is drawn from the first bag and placed unseen in the second bag. What is the probability that a ball now drawn from the second bag is black?

Sol. Let B_1, B_2, and W_1 represent, respectively, the drawing of a black ball from bag 1, a black ball from bag 2, and a white ball from bag 1. Then $B_1 \cap B_2$ represents the event that a black ball is transferred from bag 1 to the bag 2, and then a black ball is drawn from the bag 2. Likewise, $W_1 \cap B_2$ represents the event that a white ball is transferred from bag 1 to the bag 2, and then a black ball is drawn from the bag 2. So we are interested in the union of the mutually exclusive events $B_1 \cap B_2$ and $W_1 \cap B_2$. It follows that

$$P((B_1 \cap B_2) \cup (W_1 \cap B_2)) = P(B_1 \cap B_2) + P(W_1 \cap B_2)$$
$$= P(B_1)P(B_2|B_1) + P(W_1)P(W_1|B_2)$$
$$= (3/7)(6/9) + (4/7)(5/9) = 38/63.$$

Ex. Three bags carry (2 white and 3 black balls), (4 white and 5 black balls), and (6 white and 7 black balls), respectively. Find the probability of drawing a black ball from the third bag given that without noticing, a ball is transferred from first bag to the second bag, and then a ball is transferred from second bag to the third bag.

Sol. Let W_1, B_1, W_2, B_2 and B_3 be the events of drawing a while ball from first bag, a black ball from first bag, a while ball from second bag, a black ball from second bag and a black ball from the third bag respectively. Then the required probability is

$$= P(W_1 \cap W_2 \cap B_3) + P(W_1 \cap B_2 \cap B_3) + P(B_1 \cap W_2 \cap B_3) + P(B_1 \cap B_2 \cap B_3)$$
$$= P(W_1)P(W_2|W_1)P(B_3|W_1 \cap W_2) + P(W_1)P(B_2|W_1)P(B_3|W_1 \cap B_2)$$
$$+P(B_1)P(W_2|B_1)P(B_3|B_1 \cap W_2) + P(B_1)P(B_2|B_1)P(B_3|B_1 \cap B_2)$$
$$= (2/5)(5/10)(7/14) + (2/5)(5/10)(8/14) + (3/5)(4/10)(7/14) + (3/5)(6/10)(8/14)$$
$$= 378/700 = 0.54.$$

Ex. Three cards are drawn in succession, without replacement, from an ordinary deck of playing cards. Find the probability that the first card is a red ace, the second card is a 10 or a jack, and the third card is greater than 3 but less than 7.
Sol. First we define the events.
A : the first card is a red ace,
B : the second card is a 10 or a jack,
C : the third card is greater than 3 but less than 7.
$\therefore\ P(A) = \frac{2}{52}$, $P(B|A) = \frac{8}{51}$, $P(C|(A \cap B)) = \frac{12}{50}$. Thus the required probability is
$P(A \cap B \cap C) = P(A)P(B|A)P(C|(A \cap B)) = \frac{2}{52}\frac{8}{51}\frac{12}{50} = \frac{8}{5525}$.

1.2.7 Independent Events

Two events A and B are said to be independent if occurrence or non-occurrence of A does not affect the occurrence or non-occurrence of the other. Thus, if A and B are independent, then $P(A|B) = P(A)$. So $P(A \cap B) = P(A)P(B)$ is the mathematical condition for the independence of the events A and B.
Note that three events A, B, C are independent provided these are pairwise independent and $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Ex. A small town has one fire engine and one ambulance available for emergencies. The probability that the fire engine is available when needed is 0.98, and the probability that the ambulance is available when called is 0.92. In the event of an injury resulting from a burning building, find the probability that both the ambulance and the fire engine will be available, assuming they operate independently.

Sol. $(0.98)(0.92) = 0.9016$

Ex. If A and B are independent events, then show that
(i) A and B' are independent events.
(ii) A' and B' are independent events.

Sol. We need to prove $P(A \cap B') = P(A)P(B')$. We have

$A = A \cap S = A \cap (B \cup B') = (A \cap B) \cup (A \cap B')$,
where $A \cap B$ and $A \cap B'$ are mutually exclusive events. So by addition rule of probability,

$P(A) = P((A \cap B) \cup (A \cap B')) = P(A \cap B) + P(A \cap B')$

$\therefore P(A \cap B') = P(A) - P(A \cap B)$
$= P(A) - P(A)P(B)$ ($\because$ A and B are independent events.)
$= P(A)(1 - P(B))$
$= P(A)P(B')$.

Thus, A and B' are independent events.

(ii) We need to prove $P(A' \cap B') = P(A')P(B')$. We have

$P(A' \cap B') = P((A \cup B)') = 1 - P(A \cup B) = 1 - (P(A) + P(B) - P(A \cap B))$
$= 1 - (P(A) + P(B) - P(A)P(B))$ ($P(A \cap B) = P(A)P(B)$ since A and B are independent events.)
$= 1 - P(A) - P(B)(1 - P(A))$
$= (1 - P(A))(1 - P(B))$
$= P(A')P(B')$.

Thus, A' and B' are also independent events.

Ex. A problem is given to three students in a class. The probabilities of the solution from the three students are 0.5, 0.7 and 0.8 respectively. What is the probability that the problem will be solved?

Sol. $1 - (1 - 0.5)(1 - 0.7)(1 - 0.8) = 0.97$

Ex. An electrical system consists of four components as illustrated in Figure 1.5. The system works if components A and B work and either of the components C or D works. The reliability (probability of working) of each component is also shown in Figure 1.5. Find the probability that (a) the entire system works and (b) the component C does not work, given that the entire system works. Assume that the four components work independently.

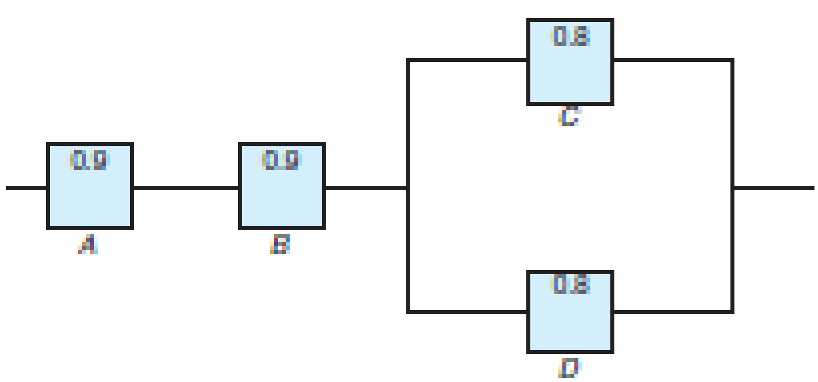

Figure 1.5:

Sol. (a) The event that the entire system works is $A \cap B \cap (C \cup D)$. Therefore its probability is,
$P(A \cap B \cap (C \cup D))$
$= P((A \cap B \cap C) \cup (A \cap B \cap D))$
$= P(A \cap B \cap C) + P(A \cap B \cap D) - P((A \cap B \cap C) \cap (A \cap B \cap D))$ (By addition rule of probability)
$= P(A \cap B \cap C) + P(A \cap B \cap D) - P(A \cap B \cap C \cap D)$
$= P(A)P(B)P(C) + P(A)P(B)P(D) - P(A)P(B)P(C)P(D)$ ($\because$ A, B, C, D all are independent events)
$= (0.9)(0.9)(0.9) + (0.9)(0.9)(0.8) - (0.9)(0.9)(0.8)(0.8)$
$= 0.7776$.

(b) The probability that the component C does not work, given that the entire system works, is given by
$P(C'|(A \cap B \cap (C \cup D)))$
$= P(A \cap B \cap (C \cup D) \cap C')/P(A \cap B \cap (C \cup D))$
$= P(A \cap B \cap ((C \cap C') \cup (D \cap C')))/P(A \cap B \cap (C \cup D))$
$= P(A \cap B \cap ((\phi) \cup (D \cap C')))/P(A \cap B \cap (C \cup D))$
$= P(A \cap B \cap D \cap C')/P(A \cap B \cap (C \cup D))$
$= P(A)P(B)P(D)P(C')/P(A \cap B \cap (C \cup D))$ ($\because$ A, B, C', D all are independent events)
$= P(A)P(B)P(D)(1 - P(C))/P(A \cap B \cap (C \cup D))$
$= (0.9)(0.9)(0.8)(1 - 0.8)/0.7776 = 0.1667$.

1.3 Theorem of Total Probability

Let B_1, B_2, ... , B_n be exhaustive and mutually exclusive events in the sample space S of a random experiment with probabilities $P(B_1)$, $P(B_2)$, ... , $P(B_n)$, respectively. Let A be any event in S, then

$$P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i).$$

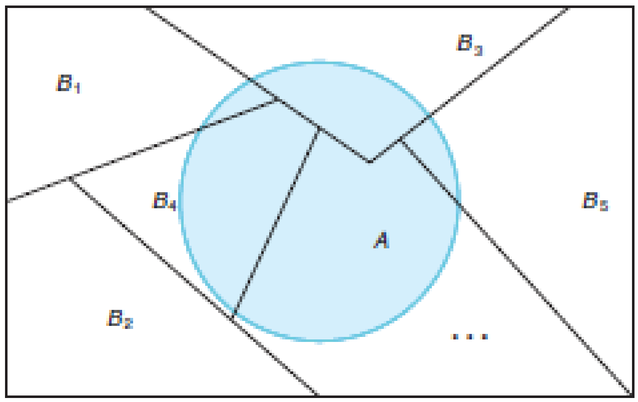

Figure 1.6: Partitioning the sample space S

Proof: Since B_1, B_2, ... , B_n are exhaustive and mutually exclusive events in the sample space S, so $S = B_1 \cup B_2 \cup\ ...\ \cup B_n$. It follows that

$$A = A \cap S = (A \cap B_1) \cup (A \cap B_2) \cup\ ...\ \cup (A \cap B_n).$$

Now B_1, B_2, ... , B_n are mutually exclusive events. Therefore, $A \cap B_1$, $A \cap B_2$, ... , $A \cap B_n$ are mutually exclusive events. So we have $P(A) = P(A \cap B_1) + P(A \cap B_2) + ... + P(A \cap B_n) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$.

Ex. In a certain assembly plant, three machines, M_1, M_2, and M_3, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected. What is the probability that it is defective?

Sol. Consider the following events:
A : the product is defective,
B_1 : the product is made by machine M_1,
B_2 : the product is made by machine M_2,
B_3 : the product is made by machine M_3.

From the given data in the statement, we have

$$P(B_1) = 0.3,\ P(B_2) = 0.45,\ P(B_3) = 0.25,$$

$$P(A|B_1) = 0.02,\ P(A|B_2) = 0.03,\ P(A|B_3) = 0.02.$$

From the total probability theorem, it follows that

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)$$
$$= (0.3)(0.02) + (0.45)(0.03) + (0.25)(0.02) = 0.0245.$$

Ex. A person has undertaken a construction job. The probabilities are 0.65 that there will be strike, 0.80 that the construction job will be completed on time if there is no strike, and 0.32 that the construction job will be completed on time if there is a strike. Determine the probability that the construction job will be completed on time.

Sol. Let A be the event that the construction job will be completed on time, and B be the event that there will be a strike. Then the required probability is given by

$$P(A) = P(B)P(A|B) + P(B')P(A|B')$$
$$= (0.65)(0.32) + (1 - 0.65)(0.8) = 0.488.$$

1.4 Bayes' Theorem

Let B_1, B_2, ... , B_n be exhaustive and mutually exclusive events in the sample space S of a random experiment with probabilities $P(B_1)$, $P(B_2)$, ... , $P(B_n)$, respectively. Let A be any event in S with $P(A) \neq 0$, then

$$P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)},$$

where $k = 1, 2, ..., n$.

Proof: From conditional probability, we have

$$P(B_k|A) = \frac{P(A \cap B_k)}{P(A)} = \frac{P(B_k)P(A|B_k)}{P(A)}.$$

So the desired result follows from the theorem of total probability.

Ex. Four units of a bulb making factory respectively produce 3%, 2%, 1% and 0.5% defective bulbs. A bulb selected at random from the entire output is found defective. Find the probability that it is produced by the fourth unit of the factory.

Sol. Consider the following events:
A : the bulb is defective,
B_1 : the bulb is made by first unit,
B_2 : the bulb is made by second unit,
B_3 : the bulb is made by third unit,
B_4 : the bulb is made by fourth unit.

From the given data in the statement, we have

$$P(B_1) = 0.25,\ P(B_2) = 0.25,\ P(B_3) = 0.25,\ P(B_4) = 0.25$$

$P(A|B_1) = 0.03,\ P(A|B_2) = 0.02,\ P(A|B_3) = 0.01,\ P(A|B_4) = 0.005.$

Then the required probability that the randomly chosen bulb from the entire output is made by the fourth unit is

$$P(B_4|A) = \frac{P(B_4)P(A|B_4)}{P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+P(B_3)P(A|B_3)+P(B_4)P(A|B_4)}$$
$$= \frac{(0.25)(0.005)}{(0.25)(0.03)+(0.25)(0.02)+(0.25)(0.01)+(0.25)(0.005)} = \frac{1}{13}.$$

Ex. Bag I contains 3 red and 4 black balls while another Bag II contains 5 red and 6 black balls. One ball is drawn at random from one of the bags and it is found to be red. Find the probability that it was drawn from Bag II.

Sol. Ans. 35/68

Ex. In a factory which manufactures bolts, machines A, B and C manufacture respectively 25%, 35% and 40% of the bolts. Of their outputs, 5, 4 and 2 percent are respectively defective bolts. A bolt is drawn at random from the product and is found to be defective. What is the probability that it is manufactured by the machine B?

Sol. Ans. 28/69

Ex. A doctor is to visit a patient. From the past experience, it is known that the probabilities that he will come by train, bus, scooter or by other means of transport are respectively 3/10, 1/5, 1/10 and 2/5. The probabilities that he will be late are 1/4, 1/3, and 1/12, if he comes by train, bus and scooter respectively, but if he comes by other means of transport, then he will not be late. When he arrives, he is late. What is the probability that he comes by train?

Sol. Let A be the event that the doctor visits the patient late and let B_1, B_2, B_3 and B_4 be the events that the doctor comes by train, bus, scooter, and other means of transport respectively. Then the required probability is $P(B_1/A)$. By the Bayes' theorem, we have

$$P(B_1|A) = \frac{P(B_1)P(A|B_1)}{P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+P(B_3)P(A|B_3)+P(B_4)P(A|B_4)}$$
$$= \frac{(3/10)(1/4)}{(3/10)(1/4)+(1/5)(1/3)+(1/10)(1/12)+(2/5)(0)} = \frac{1}{2}.$$

Ex. Given three identical boxes I, II and III, each containing two coins. In box I, both coins are gold coins, in box II, both are silver coins and in the box III, there is one gold and one silver coin. A person chooses a box at random and takes out a coin. If the coin is of gold, what is the probability that the other coin in the box is also of gold?

Sol. Let B_1, B_2 and B_3 be the events that the boxes I, II and III are chosen, respectively. Let A be the event that the coin is gold. Since the box I carries both the gold coins. So the probability that the other coin in the box is also of gold is given by $P(B_1|A)$. Ans. 2/3

Ex. A man is known to speak truth 3 out of 4 times. He rolls a die and reports that it is a six. Find the probability that it is actually a six.

Sol. Let A be the event that the man reports that six occurs in the throw of the die. Let B be the event that 6 appears on the die. Then the required probability is $P(B|A)$. By Bayes' theorem we have

$P(B|A) = \frac{P(B)P(A|B)}{P(B)P(A|B)+P(B')P(A|B')} = \frac{(1/6)(3/4)}{(1/6)(3/4)+(5/6)(1/4)} = \frac{3}{8}$.

Ex. In answering a question on a multiple choice test, a student either knows the answer or guesses. Let 3/4 be the probability that he knows the answer and 1/4 be the probability that he guesses. Assuming that a student who guesses at the answer will be correct with probability 1/4. What is the probability that the student knows the answer given that he answered it correctly?

Sol. First, we define the following events:

B_1: The student knows the answer
B_2: The student guesses the answer
A: The student correctly answers the question

Then $P(B_1) = 3/4$, $P(B_2) = 1/4$, $P(A|B_1) = 1$, $P(A|B_2) = 1/4$.
So the probability that the student knows the answer given that he answered it correctly, is given by

$P(B_1|A) = \frac{P(B_1)P(A|B_1)}{P(B_1)P(A|B_1)+P(B_1)P(A|B_1)} = \frac{12}{13}$.

1.5 Misc. Examples

Ex. Show that $P(B \cup C|A) = P(B|A) + P(C|A) - P(B \cap C|A)$.

Sol. We have
$P(B \cup C|A) = \dfrac{P((B \cup C) \cap A)}{P(A)}$
$= \dfrac{P((B \cap A) \cup (C \cap A))}{P(A)}$
$= \dfrac{P(B \cap A) + P(C \cap A) - P((B \cap A) \cap (C \cap A))}{P(A)}$
$= \dfrac{P(B \cap A) + P(C \cap A) - P(B \cap A \cap C)}{P(A)}$
$= \dfrac{P(B \cap A)}{P(A)} + \dfrac{P(C \cap A)}{P(A)} - \dfrac{P(B \cap C \cap A)}{P(A)}$
$= P(B|A) + P(C|A) - P(B \cap C|A)$.
Ex. Show that $P(C|(A \cup B)) = \dfrac{P(A)P(C|A) + P(B)P(C|B) - P(A \cap B)P(C|(A \cap B))}{P(A) + P(B) - P(A \cap B)}$.

Sol. We have
$P(C|(A \cup B))$
$= \dfrac{P(C \cap (A \cup B))}{P(A \cup B)}$
$= \dfrac{P((C \cap A) \cup (C \cap B))}{P(A \cup B)}$
$= \dfrac{P(C \cap A) + P(C \cap B) - P(C \cap A \cap B)}{P(A \cup B)}$
$= \dfrac{P(A)P(C|A) + P(B)P(C|B) - P(A \cap B)P(C|(A \cap B))}{P(A) + P(B) - P(A \cap B)}$.

Ex. Two cards are drawn from a deck of cards without replacement. Find the probability that the first card is a red card and second card is an ace.

Sol. The first red card may or may not be an ace. Let A be the event that the first red card is an ace; B be the event that the first red card is not an ace, and C be the event that the second card is an ace. Then the required probability is
$P(C \cap (A \cup B)) = P(A)P(C|A) + P(B)P(C|B) - P(A \cap B)P(C|(A \cap B))$
$= \dfrac{2}{52} \times \dfrac{3}{51} + \dfrac{24}{52} \times \dfrac{4}{51} - 0 = 0.0385.$

Ex. Two cards are drawn from a deck of cards without replacement. Find the probability that the second card is an ace given that the first card is a red card.

Sol. The first red card may or may not be an ace. Let A be the event that the first red card is an ace; B be the event that the first red card is not an ace, and C be the event that the second card is an ace. Then the required probability is
$P(C|(A \cup B)) = \dfrac{P(A)P(C|A) + P(B)P(C|B) - P(A \cap B)P(C|(A \cap B))}{P(A) + P(B) - P(A \cap B)}$
$= \dfrac{\frac{2}{52} \times \frac{3}{51} + \frac{24}{52} \times \frac{4}{51} - 0}{\frac{2}{52} + \frac{24}{52} - 0} = 0.0769.$

Remark. In the first example, we have calculated the probability of happening of the events $A \cup B$ and C. But in the second example, we have calculated the probability of happening of the event C using the given fact that the event $A \cup B$ has already happened. The certainty of event $A \cup B$ in this case gives rise to larger value of probability.

Ex. Mutually exclusive events need not be independent. For, consider two mutually exclusive events A and B such that $P(A)P(B) > 0$. Now A and B are mutually exclusive, so $A \cap B = \phi$. It follows that $P(A \cap B) = P(\phi) = 0$. Also, $P(A)P(B) > 0$. So $P(A \cap B) \neq P(A)P(B)$. It implies that A and B are not independent.

Ex. Independent events need not be mutually exclusive. For, consider two indepedent events A and B such that $P(A)P(B) > 0$. Now A and B are independent, so $P(A \cap B) = P(A)P(B)$. It follows that $P(A \cap B) > 0$. So $A \cap B \neq \phi$. It implies that A and B are not mutually exclusive.

Ex. A person has 3 different email accounts. Most of his messages, in fact 70%, come into account #1, whereas 20% come into account #2 and the remaining 10% into account #3. Of the messages into account #1, only 1% are spam, whereas the corresponding percentages for accounts #2 and #3 are 2% and 5%, respectively. What is the probability that a randomly selected message is spam?

Ans. 0.016 (Apply total probability theorem).

Ex. 1 in 1000 adults is affected with a rare disease for which a diagnostic test has been developed. The test is such that when an individual actually has the disease, a positive result will occur 99% of the time, whereas an individual without the disease will show a positive test result only 2% of the time. If a randomly selected individual is tested and the result is positive, what is the probability that the individual has the disease?

Ans. 0.047 (Apply Bayes' theorem).

Ex. Show that the condition $P(A \cap B \cap C) = P(A)P(B)P(C)$ for the independence of three events A, B and C, is necessary but not sufficient.

Sol. Suppose we roll a 8 faced die. Then the sample space has numbers 1 to 8, each with probability 1/8.

Let $A = B = \{1, 2, 3, 4\}$ and $C = \{1, 5, 6, 7\}$.

Then $P(A) = P(B) = P(C) = 1/2$ and $P(A \cap B \cap C) = P(1) = 1/8$.

However, A and B are obviously not independent. Also A and C are not independent.