Probability and statistical information are used for both operational and marketing decisions. For instance, a time series showing monthly sales is used to track the company’s growth and to set future target sales levels. Statistics such as the mean customer order size and the mean number of days a customer takes to make payments help identify the firm’s best customers as well as provide benchmarks for handling accounts receivable issues. In addition, data on monthly inventory levels are used in the analysis of operating profits and trends in product sales.

Probability is a numerical measure of the likelihood that an event will occur. Probability values are always assigned on a scale from 0 to 1. A probability near zero indicates an event is unlikely to occur; a probability near 1 indicates an event is almost certain to occur. Other probabilities between 0 and 1 represent degrees of likelihood that an event will occur.

# Counting Rules, Combinations, and Permutations

Being able to identify and count the experimental outcomes is a necessary step in assigning probabilities. We now discuss three useful counting rules.

1. Multiple-step experiments The first counting rule applies to multiple-step experiments. Consider the experiment of tossing two coins. If an experiment can be described as a sequence of k steps with n1 possible outcomes on the first step, n2 possible outcomes on the second step, and so on, then the total number of experimental outcomes is given by (n1).(n2)……(nk)
2. Combinations A second useful counting rule allows one to count the number of experimental outcomes when the experiment involves selecting n objects from a (usually larger) set of N objects. It is called the counting rule for combinations.
3. Permutations A third counting rule that is sometimes useful is the counting rule for permutations. It allows one to compute the number of experimental outcomes when n objects are to be selected from a set of N objects where the order of selection is important. The same n objects selected in a different order are considered a different experimental outcome.

# Assigning Probabilities

The three approaches most frequently used are the

1. Classical - The classical method of assigning probabilities is appropriate when all the experimental outcomes are equally likely. If n experimental outcomes are possible, a probability of 1/n is assigned to each experimental outcome. When using this approach, the two basic requirements for assigning probabilities are automatically satisfied.
2. Relative frequency [Empirical Probability] - The relative frequency method of assigning probabilities is appropriate when data are available to estimate the proportion of the time the experimental outcome will occur if the experiment is repeated a large number of times.
3. Subjective methods - The subjective method of assigning probabilities is most appropriate when one cannot realistically assume that the experimental outcomes are equally likely and when little relevant data are available. When the subjective method is used to assign probabilities to the Experimental outcomes, we may use any information available, such as our experience or intuition. After considering all available information, a probability value that expresses our degree of belief (on a scale from 0 to 1) that the experimental outcome will occur is specified. Because subjective probability expresses a person’s degree of belief, it is personal.

Regardless of the method used, two basic requirements for assigning probabilities must be met.

* The probability assigned to each experimental outcome must be between 0 and 1, inclusively.
* The sum of the probabilities for all the experimental outcomes must equal 1.

# Some Basic Relationships of Probability

Sample Space - The sample space for an experiment is the set of all experimental outcomes [all possible outcomes of the experiment]. An experimental outcome is also called a sample point to identify it as an element of the sample space. The outcomes must be exhaustive and mutually exclusive.

Random Variable - A random variable is a numerical description of the outcome of an experiment.

Event - An event is a collection or set of one or more simple events in a sample space

Complement of an Event - Given an event A, the complement of A is defined to be the event consisting of all sample points that are not in A. The complement of A is denoted by Ac.

Union - The union of events and the intersection of events. Given two events A and B, the union of A and B is defined as follows. The union of A and B is the event containing all sample points belonging to A or B or both. The union is denoted by .

Intersection - Given two events A and B, the intersection of A and B is the event containing the sample points belonging to both A and B. The intersection is denoted by . The intersection of events A and B is the event that occurs when both A and B occur. The probability of the intersection is called the joint probability.

Joint Probability

Marginal probability Marginal probabilities, computed by adding across rows or down columns, are so named because they are calculated in the margins of the table.

Addition Law - The addition law is helpful when we are interested in knowing the probability that at least one of two events occurs. That is, with events A and B we are interested in knowing the probability that event A or event B or both occur.

* Dependent Events
* Independent Events

Mutually Exclusive Events - Two events are said to be mutually exclusive if the events have no sample points in common.

# Conditional Probability

Often, the probability of an event is influenced by whether a related event already occurred. Suppose we have an event A with probability P(A). If we obtain new information and learn that a related event, denoted by B, already occurred, we will want to take advantage of this information by calculating a new probability for event A. This new probability of event A is called a conditional probability and is written P(A | B). We use the notation | to indicate that we are considering the probability of event A given the condition that event B has occurred. Hence, the notation P(A | B) reads “the probability of A given B.”

* Dependent Events
* Independent Events

Multiplication Law - Whereas the addition law of probability is used to compute the probability of a union of two events, the multiplication law is used to compute the probability of the intersection of two events. The multiplication law is based on the definition of conditional probability. Using equations above and solving for , we obtain the multiplication law.

* Dependent Events
* Independent Events

# Bayes’ Theorem

In the discussion of conditional probability, we indicated that revising probabilities when new information is obtained is an important phase of probability analysis. Often, we begin the analysis with initial or prior probability estimates for specific events of interest. Then, from sources such as a sample, a special report, or a product test, we obtain additional information about the events. Given this new information, we update the prior probability values by calculating revised probabilities, referred to as posterior probabilities. Bayes’ theorem provides a means for making these probability calculations.

# Practical Applications

1. Auditing Tax Returns - Government auditors routinely check tax returns to determine whether calculation errors were made. They also attempt to detect fraudulent returns. There are several methods that dishonest taxpayers use to evade income tax. One method is not to declare various sources of income. Auditors have several detection methods, including spending patterns. Another form of tax fraud is to invent deductions that are not real. After analyzing the returns of thousands of self-employed taxpayers, an auditor has determined that 45% of fraudulent returns contain two suspicious deductions, 28% contain one suspicious deduc- tion, and the rest no suspicious deductions. Among honest returns the rates are 11% for two deductions, 18% for one deduction, and 71% for no deductions. The auditor believes that 5% of the returns of self-employed individuals contain significant fraud. The auditor has just received a tax return for a self-employed individual that contains one suspicious expense deduction. What is the probability that this tax return contains significant fraud?
2. Mutual Funds - A mutual fund is a pool of investments made on behalf of people who share similar objectives. In most cases, a professional manager who has been educated in finance and statistics manages the fund. He or she makes decisions to buy and sell individual stocks and bonds in accordance with a specified investment philosophy. For example, there are funds that concentrate on other publicly traded mutual fund companies. Other mutual funds specialize in Internet stocks (so-called dot-coms), whereas others buy stocks of biotechnology firms. Surprisingly, most mutual funds do not outperform the market; that is, the increase in the net asset value (NAV) of the mutual fund is often less than the increase in the value of stock indexes that represent their stock markets. One reason for this is the management expense ratio (MER) which is a measure of the costs charged to the fund by the manager to cover expenses, including the salary and bonus of the managers. The MERs for most funds range from .5% to more than 4%. The ultimate success of the fund depends on the skill and knowledge of the fund manager. This raises the question, which managers do best? SOLUTION – Looking into each manager’s educational background and past performance. we would certainly like to know the probability that a fund managed by a graduate of a top20 MBA program will outperform the market. Such a probability will allow us to make an informed decision about where to invest our money. This probability is called a conditional probability because we want to know the probability that a fund will outperform the market given the condition that the manager graduated from a top-20 MBA program.

# Discrete Probability Distributions

A discrete random variable is one that can take on a countable number of values. For example, if we define x as the number of heads observed in an experiment that flips a coin 10 times, then the values of x are 0, 1, 2, . . ., 10. The variable x can assume a total of 11 values. Obviously, we counted the number of values; hence, x is discrete. For a discrete random variable x, the probability distribution is defined by a probability function, denoted by f (x). The probability function provides the probability for each value of the random variable.

### Expected value, variance & Standardard deviation of discretre random variable

The population mean is the weighted average of all of its values. The weights are the probabilities. This parameter is also called the expected value of x and is represented by E(x). The expected value, or mean, of a random variable is a measure of the central location for the random variable. The formula for the expected value of a discrete random variable x follows.

The formula for the variance of a discrete random variable follows.



## Binomial Probability Distribution

The binomial probability distribution is a discrete probability distribution that provides many applications. A binomial experiment exhibits the following four properties.

1. The experiment consists of a sequence of n identical trials.
2. Two outcomes are possible on each trial. We refer to one outcome as a success and the other outcome as a failure.
3. The probability of a success, denoted by p, does not change from trial to trial. Consequently, the probability of a failure, denoted by 1 - p, does not change from trial to trial.
4. The trials are independent.

If properties 2, 3, and 4 are present, we say the trials are generated by a Bernoulli process. If, in addition, property 1 is present, we say we have a binomial experiment.

Where:

*x* = the number of successes

*p* = the probability of a success on one trial

*n* = the number of trials

*f(x)* = the probability of x successes in n trials

Expected Value and Variance for the Binomial Distribution

## Poisson Probability Distribution

In this section we consider a discrete random variable that is often useful in estimating the number of occurrences over a specified interval of time or space (Length or Distance). For example, the random variable of interest might be the number of arrivals at a car wash in one hour, the number of repairs needed in 10 miles of highway, or the number of leaks in 100 miles of pipeline.

PROPERTIES OF A POISSON EXPERIMENT

1. The probability of an occurrence is the same for any two intervals of equal length.
2. The occurrence or non-occurrence in any interval is independent of the occurrence or non-occurrence in any other interval.

Where:

*f(x)* = the probability of x occurrences in an interval

μ = expected value or mean number of occurrences in an interval

*e* = 2.71828

## Hypergeometric Probability Distribution

If the set has two different types of classifications and we select without replacement, then we use hypergeometric probability distribution. The hypergeometric probability distribution is closely related to the binomial distribution. The two probability distributions differ in two key ways.

1. With the hypergeometric distribution, the trials are not independent [without replacement]; and
2. The probability of success changes from trial to trial.

# Continuous Probability Distributions

A continuous random variable is one that can assume an uncountable number of values. Because this type of random variable is so different from a discrete variable, we need to treat it completely differently. First, we cannot list the possible values because there is an infinite number of them. Second, because there is an infinite number of values, the probability of each individual value is virtually 0. Consequently, we can determine the probability of only a range of values.

the probability density function does not directly provide probabilities. However, the area under the graph of f (x) corresponding to a given interval does provide the probability that the continuous random variable x assumes a value in that interval. Because the area under the graph of f (x) at any particular point is zero, one of the implications of the definition of probability for continuous random variables is that the probability of any particular value of the random variable is zero.

## Uniform/Rectangular probability Distribution

probability can be computed as the area under the graph of f (x), area of a rectangle is simply the width multiplied by the height. height equal to the value of the probability density function f (x)

Where

## Normal Probability Distribution

The most important probability distribution for describing a continuous random variable is the normal probability distribution. the normal distribution is illustrated by the bell-shaped normal curve.

NORMAL PROBABILITY DENSITY FUNCTION

Where:

µ = mean

σ = standard deviation

π = 3.14159

e = 2.71828

We make several observations about the characteristics of the normal distribution.

1. The entire family of normal distributions is differentiated by two parameters: the mean μ and the standard deviation σ.
2. The highest point on the normal curve is at the mean, which is also the median and mode of the distribution.
3. The mean of the distribution can be any numerical value: negative, zero, or positive.
4. The normal distribution is symmetric, with the shape of the normal curve to the left of the mean a mirror image of the shape of the normal curve to the right of the mean. The tails of the normal curve extend to infinity in both directions and theoretically never touch the horizontal axis. Because it is symmetric, the normal distribution is not skewed; its skewness measure is zero.
5. The standard deviation determines how flat and wide the normal curve is. Larger values of the standard deviation result in wider, flatter curves, showing more variability in the data. Two normal distributions with the same mean but with different standard deviations are shown here.
6. Probabilities for the normal random variable are given by areas under the normal curve. The total area under the curve for the normal distribution is 1. Because the distribution is symmetric, the area under the curve to the left of the mean is .50 and the area under the curve to the right of the mean is .50.
7. The percentage of values in some commonly used intervals are
8. 68.3% of the values of a normal random variable are within plus or minus one standard deviation of its mean.
9. 95.4% of the values of a normal random variable are within plus or minus two standard deviations of its mean.
10. 99.7% of the values of a normal random variable are within plus or minus three standard deviations of its mean.

#### Standard Normal Probability Distribution

A random variable that has a normal distribution with a mean of zero and a standard deviation of one is said to have a standard normal probability distribution. The letter z is commonly used to designate this particular normal random variable. It has the same general appearance as other normal distributions, but with the special properties of μ = 0 and σ = 1. Because μ = 0 and σ = 1, the formula for the standard normal probability density function is a simpler

STANDARD NORMAL PROBABILITY DENSITY FUNCTION

As with other continuous random variables, probability calculations with any normal distribution are made by computing areas under the graph of the probability density function. Thus, to find the probability that a normal random variable is within any specific interval, we must compute the area under the normal curve over that interval.

## Exponential Probability Distribution