**Concept of Probability and Probability Distributions**

An **experiment** is any process of observation that has an uncertain outcome. (Or a process that produces a single outcome whose result cannot be predicted with certainty is called an **experiment**.)

For instance, a very simple experiment might involve flipping a coin and the two possible outcomes are head and tail.

The collection of all outcomes that can result from a selection, decision or experiment is known as **sample space**.

A collection of experimental (or sample space) outcomes is called an **event**.

Two events are **mutually exclusive** if the occurrence of one event precludes the occurrence of the other event.

Two events are **independent** if the occurrence of one event in no way influences the probability of the occurrence of the other event.

Two evens are **dependent** if the occurrence of one event impacts the probability of the other event occurring.

Outcomes of trial are said to be **equally likely** if taking into consideration all the relevant evidences, there is no reason to expect one in preference to the others.

The concept of probability is one that we all encounter in everyday life. When a weather forecaster states that the chance of rain is 70%, he or she is making a probability statement. When a sports commentator states that the odds against the Miami Heat winning the NBA Championship are 3 to 1, he or she is making a probability statement. The concept of probability is quite intuitive. However, the rules of probability are not always as intuitive or easy to master. The most important of these rules are:

* A probability is a number between 0 and 1 that measures the likelihood that some event will occur. An event with probability 0 cannot occur, whereas an event with probability 1 is certain to occur. An event with probability greater than 0 and less than 1 involves uncertainty. The closer its probability is to 1, the more likely it is to occur.

Probabilities are sometimes expressed as percentages or odds. However, these can easily be converted to probabilities on a 0-to-1 scale. If the chance of rain is 70%, then the probability of rain is 0.7. Similarly, if the odds against the Heat winning are 3 to 1, then the probability of Heat winning is  (or 0.25).

* **Rule of complements:** It’s the simplest probability rule. If A is any event, then the complement of A, denoted by , is the event that A does not occur. If the probability of A is , then the probability of its complement, , is given by

****

Equivalently, the probability of an event and the probability of its complement sum to 1.

* All possible outcomes associated with an experiment form the sample space. Therefore, the sum of the probabilities of all possible outcomes is 1, that is,



Events are **mutually exclusive** if at most one of them can occur. That is, if one of them occurs, then none of the others can occur.

**Exhaustive events** mean that they exhaust all possibilities. The total number of possible outcomes of a random experiment is known as the exhaustive events.

* **Addition Rule:** The probability of an event  is equal to the sum of the probabilities of the individual outcomes forming . For example, if 

Note that, the individual outcomes are mutually exclusive.

* **Conditional probability** is the probability that an event will occur given that some other event has already happened. Then, for any two events  and , the probability of event  given event  has occurred is



* The multiplication rule for  is



The conditional probability formula and the multiplication rule are both equivalent. The one you use depends on which probabilities you know and which you want to calculate.

* For independent events  and , multiplication rule becomes



The joint probability of two independent events is simply the product of the probabilities of the two events. This rule is the primary way that you can determine whether any two events are independent. If the product of the probabilities of the two events equals the joint probability, then the events are independent.

**Methods of assigning probabilities:**

There are three common ways to assign probability to outcomes: classical probability assessment, relative frequency of occurrence, and subjective probability assessment.

* **Classical probability or a priori probability assessment:** The method of determining probability based on the ratio of the number of ways an outcome or event of interest can occur to the number of ways any outcome or event can occur when the individual outcomes are equally likely.

****

* **Relative frequency assessment:** The method that defines probability as the number of times an event occurs divided by the total number of times an experiment is performed in a large number of trials.

****

* When we use experience, intuitive judgment, or expertise to assess a probability, we call this the **subjective method of assigning probability**. A subjective probability is a measure of a personal conviction that an event will occur. Therefore, in this instance, probability represents a person’s belief that an event will occur.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Bayes’ Theorem**

As decision makers, you will often encounter situations that require you to assess probabilities for events of interest. Your assessment may be based on relative frequency or subjectivity. However, you may then come across new information that causes you to revise the probability assessment. For example, a human resource manager who has interviewed a person for a sales job might assess a low probability that the person will succeed in sales. However, after seeing the person’s very high score on the company’s sales aptitude test, the manager might revise her assessment upward. A medical doctor might assign an 80% chance that a patient has a particular disease. However, after seeing positive results from a lab test, he might increase his assessment to 95%.

In these situations, you will need a way to formally incorporate the new information. One very useful tool for doing this is called Bayes’ Theorem, which is named for the Thomas Bayes, who developed the special application of conditional probability in the 1700s.

**Bayes’ Theorem**: If  are mutually disjoint events with , then for any arbitrary event A which is a subset of  such that , we have



\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Probability Distributions**

A **probability model** is a mathematical representation of a random phenomenon.

Random variable is a type of random phenomenon. A **random variable** is a variable whose value is numeric and is determined by the outcome of an experiment.

A **discrete** random variable is a random variable that can only assume values that are finite or countable.

Examples:

1. The number, , of the next three customers entering a store who will make a purchase. Here  could be 0, 1, 2, or 3. In this case the random variable is countable and finite.
2. The number, , of major fires in a large city in the next two months. Here  could be 0, 1, 2, 3, and so forth (there is no definite maximum number of fires). In this case the random variable is countable and infinite.

A random variable is said to be **continuous** if it can take any value in an interval. For example, the exact time it takes a trainee to perform a job task may be any value between two points, say 1 minute to 10 minutes. If  is the time required, then x is continuous because, if measured precisely enough, the possible values, , can be any value in the interval 1 to 10 minutes.

A continuous random variable is generally defined by measuring, which is contrasted with a discrete random variable, whose value is typically determined by counting.

The probability model describing a random variable is called a **probability distribution** and consists of

1. a specification of the possible values of the random variable and
2. a table, graph, or formula that can be used to calculate probabilities concerning the values that the random variable might equal.

Every probability distribution must satisfy each of the following three requirements.

1. There is a numerical random variable  and its values are associated with corresponding probabilities.
2.  where  assumes all possible values. (The sum of all probabilities must be 1, but sums such as 0.999 or 1.001 are acceptable because they result from rounding errors.)
3.  for every individual value of the random variable . (That is, each probability value must be between 0 and 1 inclusive.)

The second requirement comes from the simple fact that the random variable  represents all possible events in the entire sample space, so we are certain that one of the events will occur. The third requirement comes from the basic principle that any probability value must be 0 or 1 or a value between 0 and 1.

**Probability function** is a rule that assigns probabilities to the values of the random variables. A probability function can be as simple as a list that pairs the values of a random variable with their probabilities. However, a probability function is most often expressed in formula form.

The probability distribution for a discrete random variable is called **discrete probability distribution** and the probability functions associated with discrete probability distributions are known as **probability mass functions** (pmf).

The probability distributions associated with continuous random variables are called **continuous probability distributions** and corresponding probability functions are called **probability density functions** (pdf).

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Moments**

Moment is a term generally used in physics or mechanics and provides us a measure of the turning or the rotation effect of a force about some point. However, the term moment as used in physics has nothing to do with the moment used in Statistics, the only analogy being that in Statistics we talk of moment of a random variable about some point and these moments are used to describe the various characteristics of a probability distribution viz., central tendency, dispersion, skewness and kurtosis. Moments can also be calculated about any arbitrary point . Such moments are called raw moments.

The moment about origin is



The first moment about origin gives mean. That is,



The  moment about mean is



The moments about mean are called the central moments.

The first moment about mean is always zero because the algebraic sum of deviations of a given set of observations from their mean is zero. So,



The second moment about mean gives the variance of the distribution.



The third and fourth moments about mean in terms of moments about origin are



The first four moments enable us to have a fairly good idea about the nature and the form of the given probability distribution, in practice we generally compute only the first four moments and not the higher moments.

For a symmetrical distribution, all the odd order moments about mean vanish. That is, in case of a symmetrical distribution, if the deviations of the given observations from their arithmetic mean are raised to any odd power, the sum of positive deviations equals the sum of negative deviations and accordingly the overall sum is zero. That is,



The moments about mean are invariant under change of origin but not of scale. Let us change the origin and scale in the variable  to obtain a new variable  as defined below:



Then we can find that



Hence,  moment of the variable  about its mean is equal to  times the  moment of the variable  about its mean. The result does not depend on .

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Moment Generating Functions

The moment generating function of a random variable X (about origin) having the probability function  is given by:



In general, the moment generating function  about the point  is defined as

