# Chapter-1.1

## 1) Discrete and Continuous Probability Distributions:

a. In statistics we come across dozens of different types of probability distributions like Binomial, Poisson, Exponential, etc. All of these distributions can be classified as either a discrete or continuous probability distributions.

b. For example, let's say you have a choice of playing two games at fair chance.

   i. **Game 1:** Roll a die. If you roll a six, you win a prize.

      **Game 2:** Guess the weight of the man. If you guess within 10 pounds, you win a prize.

      One of these games is a discrete probability distribution and one is a continuous probability distribution. Which is which?

   ii. It is quite visible that Game 1 is discrete probability distribution as you can role a 1,2,3,4,5, or 6. All of the die rolls have an equal chance of being rolled (one out of six, or 1/6).

   iii. For the guess the weight game, you could guess that the man weighs 150 lbs. Or 210 pounds. Or 185.5 pounds. Or any fraction of a pound (172.566 pounds). Even if you stick to, say, between 150 and 200 pounds, the possibilities are endless:

      1. 160.1 lbs, 160.11 lbs, 160.111 lbs, 160.1111 lbs.

   iv. In reality, you probably wouldn't guess 160.111111 lbs…that seems a little ridiculous. But it doesn't change the fact that you *could* (if you wanted to), so that's why it's a **continuous probability distribution**.

c. Following are examples of discrete probability distributions:

      1. Binomial, Poisson distributions.

d. Following are examples of continuous probability distributions:

      1. Uniform, Exponential, Normal distributions.

e. We say that X has a discrete distribution if X is a discrete random variable. In particular, for some finite or countable set of values x1, x2, … we have $P(X = x_i) > 0$, i = 1,2, … and $\sum_i P(X = x_i) = 1$.We define the probability mass function (pmf) by f(x) = P(X = x).

   i. Discrete random variables can only take isolated values.

f. A random variable X is said to have a continuous distribution if X is a continuous random variable for which there exists a positive function f with total integral 1, such that for all a, b. The function f is called the probability density function (pdf) of X.

   i. Continuous random variables can take values in an interval.

ii. $P(a < X \leq b) = F(b) - F(a) = \int_a^b f(u)d(u)$

iii. However, it is important to realise that f(x) is not a probability -- is a probability density. In particular, if X is a continuous random variable, then $P(X = x) = 0$, for all x. As we cannot say the weight is exact 70kg, it can be 69.99999kg and so on.

## 2) Binomial Distribution:

a. Consider a sequence of n coin tosses. If X is the random variable which counts the total number of heads and the probability of "head" is p then we say X has a binomial distribution with parameters n and p and write $X \sim Bin(n, p)$. The probability mass function X is given by:

i. $f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ , x = 0, 1, …, n

b. The expectation is EX = np. To prove this, simple evaluate the sum:

i. $\underline{\sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}}$

Also we will prove that the expectation of such a sum is the sum of the expectation, therefore, $EX = E(X_1 + \ldots + X_n) = EX_1 + \ldots + EX_n = p + \ldots + p$ [n times] = np.

c. The variance of X is Var(X) = np(1- p). This is proved in a similar way to the expectation: $Var(X) = Var(X_1 + \ldots + X_n) = Var(X_1) + \ldots + Var(X_n) = p(1-p) + \ldots + p(1-p)$ [n times] = np(1-p).

## 3) Poisson Distribution:

a. A random variable X for which:

i. $P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}, x = 0,1,2,\ldots$

b. (for fixed $\lambda > 0$) is said to have a Poisson distribution. We write $X \sim Poi(\lambda)$. The Poisson distribution is used in many probability models and may be viewed as the "limit" of the Bin(n, μ/n) for large n in the following sense: Consider a coin tossing experiment where we toss a coin n times with success probability $\lambda$/n.

c. Let X be the number of successes. Then, as we have seen:

i. $X \sim Bin(n, \lambda/n)$.

d. In other words,

i. $P(X = k) = \binom{n}{k} . \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$

1. $= \frac{\lambda^k}{k!} . \frac{n \, X \, n-1 \, X \, n-2 \ldots X \, n-k+1}{n \, X \, n \, X \ldots . X \, n} . \left(1 - \frac{\lambda}{n}\right)^n . \left(1 - \frac{\lambda}{n}\right)^{-k}$

ii. As $n \to \infty$, the second and fourth factors go to 1, and the third factor goes to $e^{-\lambda}$ (this is one of the defining properties of the exponential function). Hence, we have:

1. $\lim_{n\to\infty} P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}$

which shows that the Poisson distribution is a limiting case of the binomial one.

e. It follows that the expectation is EX = $\lambda$. The intuitive explanation is that the mean number of successes of the corresponding coin flip experiment is np = n($\lambda/n$) = $\lambda$.

f. The above argument suggests that the variance should be n($\lambda/n)(1-\lambda/n)$ = $\lambda$. This is indeed the case, as

i. Var(X) = $\lambda^2 + \lambda - \lambda^2 = \lambda$.

ii. Thus for Poisson distribution expectation and variance are same.

## 4) Uniform Distribution:

a. We say that a random variable X has a uniform distribution on the interval [a, b], if it has density function f, given by:

i. $f(x) = \frac{1}{b-a}, a \le x \le b$

b. We write X~U[a, b]. X can model a randomly chosen point from the interval [a, b], where each choice is equally likely.

c. Expectation:

i. $EX = \int_a^b \frac{x}{b-a}dx = \frac{1}{b-a}\left[\frac{b^2-a^2}{2}\right] = \frac{a+b}{2}$

d. Variance:

i. $Var(X) = EX^2 - (EX)^2 = \int_a^b \frac{x^2}{b-a}dx - \left(\frac{a+b}{2}\right)^2 = \cdots = \frac{(b-a)^2}{12}$

## 5) Exponential Distribution:

a. A random variable X with probability density function f, given by

i. $f(x) = \lambda e^{-\lambda x}, x \ge 0$, is said to have an exponential distribution with parameter $\lambda$. We write X~ Exp($\lambda$). The exponential distribution can be viewed as a continuous version of the geometric distribution.

b. Expectation:

i. EX = $\frac{1}{\lambda}$

c. Variance:

i. $Var(X) = EX^2 - (EX)^2 = \frac{1}{\lambda^2}$

d. For example, when X denotes the lifetime of a machine, then given the fact that the machine is alive at time s, the remaining lifetime of the machine, i.e. X - s, has the same exponential distribution as a completely new machine. In other words, the machine has no memory of its age and does not “deteriorate" (although it will break down eventually).

e. It is not too difficult to prove that the exponential distribution is the only continuous (positive) distribution with the memoryless property.

6) Normal or Gaussian Distribution:

a. The normal (or Gaussian) distribution is the most important distribution in the study of statistics. We say that a random variable has a normal distribution with parameters μ and $\sigma^2$ if its density function f is given by:

   i. $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2}, x \in R$

b. We write $X \sim N(\mu, \sigma^2)$. The parameters μ and $\sigma^2$ to be expectation and variance of the distribution, respectively. If $\mu = 0$ and $\sigma = 1$ then:

   i. $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\cdot(x)^2}$ and the distribution is known as a standard normal distribtution.

c. $If\ X \sim N(\mu,\ \sigma^2), and\ Z = \frac{(X-\mu)}{\sigma}$. Then,

   i. $\frac{(X-\mu)}{\sigma} \sim N(0,1)$.

   ii. Thus, by subtracting the mean and dividing by the standard deviation we obtain a standard normal distribution. This procedure is called standardisation.

   iii. Thus, Z has a standard normal distribution.

   iv. Standardisation enables us to express the cdf of any normal distribution in terms of the cdf of the standard normal distribution.

d. A trivial rewriting of the standardisation formula gives the following important result: If $X \sim N(\mu, \sigma^2)$, then:

   i. $X = \mu + \sigma Z, with\ Z \sim N(0,1)$

   ii. In other words, any Gaussian (normal) random variable can be viewed as a so-called affine (linear + constant) transformation of a standard normal random variable.

   iii. $EX = \mu$. This is because the pdf is symmetric around μ.

   iv. $Var(X) = \sigma^2$.

7) For this all distribution also refer to cdf and pdf for it.

- A data structure is a particular way of storing and organizing data in a computer so that it can be used efficiently.
- Conceptual (high-level, semantic)
  - it is working with concepts that are close enough that most users think about data
  - called entity or object-based data model too
- Physical (low-level, internal)

- it is wotking with concepts that describe the way the data is stored on your computer

abstract data types (homogeneous and heterogeneous, static and dynamic, structure, operations).:

*Homogeneous data*: Homogeneous data structures are those data structures that contain only similar type of data e.g. like a data structure containing only integer or float values. The simplest example of such type of data structures is an Array.

*Heterogeneous Data*: Heterogeneous Data Structures are those data structures that contains a variety or dissimilar type of data, for e.g. a data structure that can contain various data of different data types like integer, float and character. The examples of such data structures include structures, union etc.

A lot of people define static typing and dynamic typing with respect to the point at which the variable types are checked. Using this analogy, static typed languages are those in which type checking is done at compile-time, whereas dynamic typed languages are those in which type checking is done at run-time.

Elementary data structures: lists, stacks, queues. Sets, multisets, arrays. The representation of trees, tree traversal, deletion and insertion.

List: http://www.yamamo10.jp/yamamoto/lecture/2007/2E/test_2/html/img4.png

- Use pointers to indicate order.

- There is no need to store data in contiguous areas of memory. There is no need to store data in a contiguous area of memory, nor is there a need to store values in the order of the data.

Stacks:

A stack is a data structure that allows the last data to be retrieved first. It is a data structure as shown in http://www.yamamo10.jp/yamamoto/lecture/2007/2E/test_2/html/img32.png

Since it is a data structure, it can store data and retrieve it. The characteristic of the stack is that the data stored last is retrieved first. The data to be retrieved is the latest data stored, and since the last data is retrieved first, it is called LIFO (last in first out). It is called LIFO (last in first out).

Stacking data on the stack - storing data - is called push, and removing data from the stack is called pup.

Queue: A queue is a structure that is also called a queue. It means a queue waiting for its turn to be lined up at the window, and the data structure is as shown in . While in a stack, data is inserted and removed only from one side of the column, in a queue, data is inserted and removed from both sides of the column. One side is used for adding data and the other side is used for retrieving data. In a queue, the first data to be inserted is the first to be retrieved. The data to be retrieved is the oldest data stored, and is called FIFO (first in first out) because the first data is retrieved first. As with stacks, it is not allowed to retrieve data in the middle of a stack.