STA201

Introduction to Random Variables and Mathematical Expectation

13.1 – Introduction to Random Variables

13.1.1 - Random Variables

What are Random Variables?

A random variable is a variable which takes specified values with specified probability. A random variable is a variable that takes on numerical values as a result of a random experiment or measurement, and associates a probability with each possible outcome. Mathematically, a random variable is a real-valued function defined over a sample space. Random variables are generally denoted by uppercase letters, such as X, Y, Z etc.

Example: Consider the experiment of flipping two fair coins. The possible outcomes are: S = {HH, HT, TH, TT}. Let, X = The number of heads. So, X = {0, 1, 2}
And the probabilities associated with each value of X can be represented by the following table:

X = x	0	1	2
P(X = x)	1/4	2/4	1/4

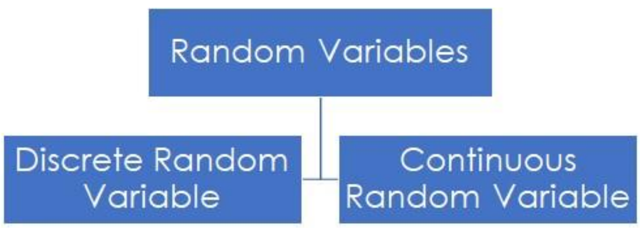
Goals scored by Messi in different matches:

Goal=1 Match-01	Goal=1 Match-02	Goal=2 Match-03
Goal=0 Match-04	Goal=1 Match-05	Goal=3 Match-06
Goal=4 Match-07	Goal=2 Match-08	Goal=05 Match-09

And the probabilities associated with each value of X can be represented by the following table:

X = x	0	1	2	3	4	5
P(X = x)	1/9	3/9	2/9	1/9	1/9	1/9

Types of Random Variables:



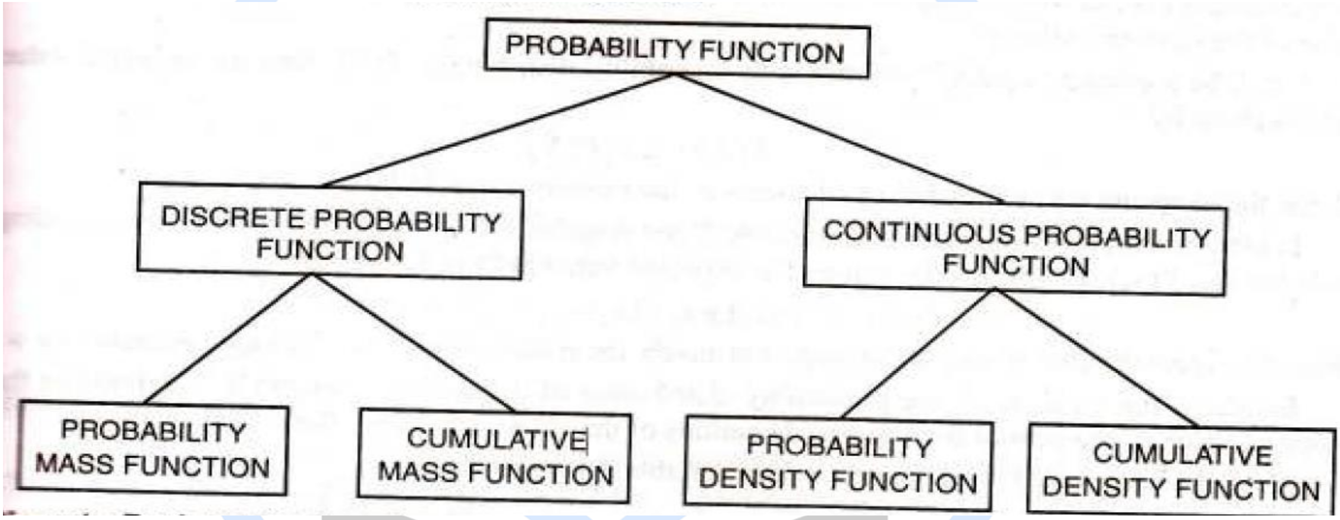
Discrete Random Variables: A discrete random variable is a random variable whose possible values either constitutes a finite set of values or an infinite sequence of numbers that is a countably infinite set of numbers. Example:

- X = The number of cars crossing an intersection every hour
- X = The number of phone calls received per day at a call center.
- X = Number of employee hire by a company.

Continuous Random Variable: A random variable is said to be continuous whose possible values consists of either all values of a small interval on real number line or all numbers in a disjoint union of such intervals (e.g. [0, 5] U [10, 15]). Continuous random variables can represent any value within a specified range or interval and can take on an infinite number of possible values. Example:

- X = The time taken to serve a customer at a call center
- X = The daily temperature at noon

13.1.2 - Probability Functions of Random Variables:



Probability Mass Function (PMF)

The probability distribution of a discrete random variable is known as discrete probability distribution.

If X is a discrete random variable with possible values x_1, x_2, … , x_n, where each value has a corresponding probability $P(X = x_i)$; $i = 1, 2, \ldots, n$, the probability mass function P(x) of X is defined by

$$P(x_i) = \begin{cases} P(X = x_i)\ ;\ \text{if } X = x_i\ , i = 1,2,\ldots,n \\ 0\ ;\ Otherwise \end{cases}$$

And has following properties:

1. $0 \le P(x) \le 1$ for all x_i
2. $\sum_{i=1}^{n} P(x) = 1$

Example 1:

Consider the experiment of flipping two fair coins.

Let X = The number of heads

So, X = {0, 1, 2}

$$P(x_i) = \begin{cases} P(X = xi) \text{ ; if } X = xi \text{ , } i = 1,2,\dots,n \\ 0 \text{ ; } Otherwise \end{cases}$$

x_i	0	1	2
P (x_i)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Example 2: Let X be a random variable with probability function defined as follows:

x_i	2	4	6	8
P(x_i)	2/10	1/10	4/10	3/10

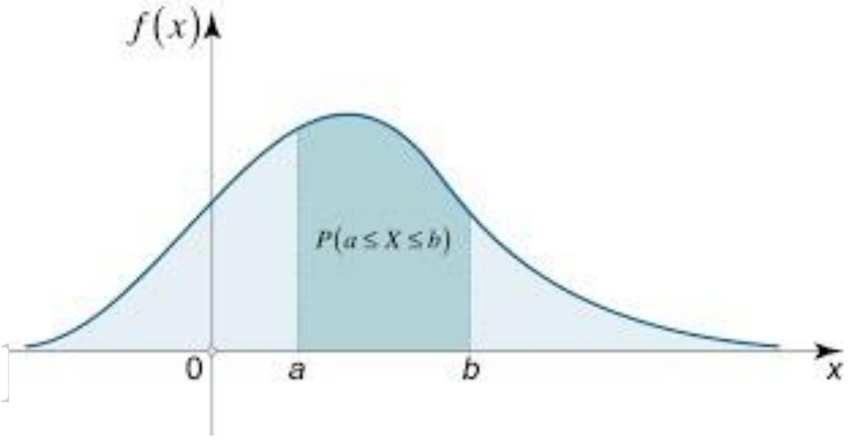
Probability Density Function (PDF)

The probability distribution of a continuous random variable is known as continuous probability distribution

If X is a continuous random variable, the probability density function $f(x)$ of X is a function such that for any two numbers a and b with a ≤ b

$$P[a \le x \le b] = \int_a^b f(x)\,dx$$

That is, the probability that X takes on a value in the interval [a, b] is equivalent to the area below the graph of $f(X)$ between the interval [a, b]. The graph of $f(X)$ is often referred to as the density curve.



And a valid PDF f(x) has the following properties

1. $0 \le f(x) \le 1$

2. $\int_{-\infty}^{\infty} f(x) = 1$

Note: The value of P(X) for any point value of X, say X=k, will always be 0. That is,

$$P(X = k) = 0; \quad k \in X$$

Example: The probability density function of a random variable X is defined as

$$f(x) = \begin{cases} x; & 0 \le x < 1 \\ 2 - x; & 1 \le x < 2 \\ 0; & x \ge 2 \end{cases}$$

Find $P[0.5 \le x \le 1.5]$

Sol: $P(0.5 \le x \le 1.5) = \int_{0.5}^{1.5} f(x)\,dx$

$$= \int_{0.5}^{1} x\,dx + \int_{1}^{1.5} (2 - x)\,dx$$

$$= \left[\frac{x^2}{2}\right]_{0.5}^{1} + \left[2x - \frac{x^2}{2}\right]_{1}^{1.5}$$

$$= \frac{3}{4}$$

13.2 – Mathematical Expectation of Random Variables

13.2.1 - Expectation of Discrete Random Variables:

Mathematical Expectation

Let X be a random variable with probability function $\mathrm{P}(x)$ (if X is discrete), or density function $f(x)$ (if X is continuous). Let $g(x)$ be a function of the random variable X. Then, the mathematical expectation of the random variable $g(x)$ is defined by

$$E[g(x)] = \begin{cases} \sum g(x) \cdot \mathrm{P}(x); & if\ X\ is\ discrete \\ \int g(x) \cdot f(x)\,dx; & if\ X\ is\ continuous \end{cases}$$

$E[g(x)]$ is also known as the expected value of $g(x)$, or the mean of the distribution of $g(x)$.

Expectation of Discrete Random Variable

Let X be a discrete random variable which can take a finite or infinite sequence of possible values x_1, x_2, … , x_n, … with corresponding probabilities $\mathrm{P}(x_1)$, $\mathrm{P}(x_2)$, … , $\mathrm{P}(x_n)$, … ; then the mathematical expectation of the random variable X, denoted by μ is defined as

$$\mu = E[X] = \sum_{i=1}^{n} x_i \mathrm{P}(x_i)\ ; \quad if\ X\ is\ finite$$

Example 1:

Let X e a random variable with probability function defined as follows:

x	2	4	6	8
$\mathrm{P}(x)$	2/10	1/10	4/10	3/10

What is the expected value of X?

Solution:

$E(x) = \sum x_i \cdot \mathrm{P}(x_i)$

$= (2 \times 2/10) + (4 \times 1/10) + (6 \times 4/10) + (8 \times 3/10)$

= 5.6

Example 2:

Imagine a game in which, on any play, a player has a 20% chance of winning Tk. 30 and an 80% chance of losing Tk. 10. What is the expected gain/loss of the player in the long run?

Solution: Let X = the gain on a play

x	30	-10
P(x)	0.2	0.8

$E(x) = \sum x_i \cdot P(x_i)$

$= (30 \times 0.2) + (-10 \times 0.8)$

$= -2$

Example 3:

If the random variable X is the top face of a tossed, fair, six-sided die, what is the expected value of X?

Solution: $X = \{1, 2, 3, 4, 5, 6\}$

P(x)= 1/6; *for x = 1, 2, 3, 4, 5, 6*

x	1	2	3	4	5	6
P(x)	1/6	1/6	1/6	1/6	1/6	1/6

$E(x) = \sum x_i \cdot P(x_i)$

$= (1\times 1/6) + (2\times 1/6) + (3\times 1/6) + (4\times 1/6) + (5\times 1/6) + (6\times 1/6)$

$= 3.5$

-

13.2.2 Expectation of Continuous Random Variables:

Expectation of Continuous Random Variables

Let X be a continuous random variable with probability density function $f(x)$; then the mathematical expectation of the random variable X, denoted by μ is defined as

$$\mu = E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

Example 1:

Suppose X is a continuous random variable with probability density function

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & otherwise \end{cases}$$

What is the expected value of X?

Solution:

$$E(x) = \int_0^1 x . f(x)\, dx$$

$$= \int_0^1 x . 2x\, dx$$

$$= \int_0^1 2x^2\, dx$$

$$= \left[\frac{2x^3}{3}\right]_0^1$$

$$= \frac{2}{3}$$

Example 2: The probability density function of a random variable X is defined as

What is the expected value of X?

$$f(x) = \begin{cases} x; & 0 \le x < 1 \\ 2 - x; & 1 \le x < 2 \\ 0; & x \ge 2 \end{cases}$$

Solution:

$$E(x) = \int_{-\infty}^{\infty} x . f(x)\, dx$$

$$= \int_0^2 x . f(x)\, dx$$

$$= \int_0^1 x . f(x)\, dx + \int_1^2 x . f(x)\, dx$$

$$= = \int_0^1 x \times x dx + \int_1^2 x \times (2 - x) dx$$

-

$$= \int_0^1 x^2\,dx + \int_1^2 2x - x^2\,dx$$
$$= \left[\frac{x^3}{3}\right]_0^1 + \left[x^2 - \frac{x^3}{3}\right]_1^2$$
$$= 1$$

13.2.3 Properties of Mathematical Expectation:

Expectation of Functions of a Random Variable

Let X be a random variable with probability function $f(x)$. Let $g(x)$ be a function of the random variable X. Then, the mathematical expectation of the function $g(x)$ is defined by

$$E[g(x)] = \begin{cases} \sum g(x) \cdot f(x); & if\ X\ is\ discrete \\ \int g(x) \cdot f(x)\,dx; & if\ X\ is\ continuous \end{cases}$$

For example, for a random variable X with probability function $f(x)$, the expected value of X^2 is

$$E[X^2] = \begin{cases} \sum X^2 \cdot f(x); & if\ X\ is\ discrete \\ \int X^2 \cdot f(x)\,dx; & if\ X\ is\ continuous \end{cases}$$

Linearity of Expectation

Let X and Y be two random variables, and let c be a constant.

Consequently, $E[X]$ and $E[Y]$ are the expected values of X and Y respectively Then,

the following properties are true:

- $E[c] = c$
- $E[cX] = c\,E[X]$
- $E[X + c] = E[X] + c$
- $E[X + Y] = E[X] + E[Y]$
- $E[X - Y] = E[X] - E[Y]$

Multiplicity of Expectation

Let X and Y be two independent random variables, and $E[X]$ and $E[Y]$ are the expected values of X and Y respectively. Then,

$$E[XY] = E[X] \cdot E[Y]$$

13.2.4 Variance of Random Variables

Variance

$$\sigma^2 = Var(X) = E[X - E(X)]^2 = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2$$

-

Standard Deviation

$$\sigma = SD(X) = \sqrt[2]{Var(X)}$$

Example:

You want to open a new Café. After your market research, you found that 20% of similar cafés make a monthly loss of Tk. 50,000, 30% of them make no profit or loss, 40% make a profit of Tk. 50,000, and 10% of them make a profit of Tk. 150, 000.

a) What is your expected profit if you decide to open a new Café?

b) What is the standard deviation in your profit amount?

c) If your fixed cost increases by Tk. 10,000, what will be your new expected profit?

Solution: Let X = profit amount

$X = x$	-50,000	0	50,000	150,000
$f(x)$	0.2	0.3	0.4	0.1

a) $E(X) = (-50{,}000 \times 0.2) + (0 \times 0.3) + (50{,}000 \times 0.4) + (1{,}50{,}000 \times 0.1) = 25000$

b) $\sigma^2 = Var(X) = E(X^2) - [E(X)]^2$

$E[X^2] = (-50{,}000^2 \times 0.2) + (0^2 \times 0.3) + (50{,}000^2 \times 0.4) + (1{,}50{,}000^2 \times 0.1)$
$= 3750000000$

$\therefore Var(X) = \sigma^2 = 3750000000 - (25000)^2 = 3125000000$

$\therefore SD(X) = \sigma = \sqrt{3125000000} = 55901.69944$

c) $E(X - 10000) = E(X) - E(10000) = E(X) - 10000 = 25000 - 10000 = 15000$

Properties of Variance

Let X and Y be two independent random variables, and $Var[X]$ and $Var[Y]$ are the variances of X and Y respectively. Let c be a constant.

Then, the following properties are true:

- $Var(c) = 0$
- $Var(c\,X) = c^2\,Var(x)$
- $Var(X + c) = Var(x)$
- $Var(X + Y) = Var(X) + Var(Y)$
- $Var(X - Y) = Var(X) + Var(Y)$

Practice Problems

Probability & Statistics for Engineering and the Sciences (Devore)

Random Variables (Basic Concept)

Page 95-96: 7

Discrete Random Variables

Page 104-105: 11(a, c), 13, 15(a, b), 17, 19, 27

Page 113-114: 29, 35, 37, 39

Continuous Random Variables

Page 142-143: 3, 5, 7

Page 150-152: 15(b, e, f), 21, 23