PES UNIVERSITY

ENGINEERING MATHEMATICS - II

Discrete Probability Distribution

Discrete probability distribution

Discrete Random Variable

Probability Mass Function (PMF)

Cumulative Distribution Function (CDF)

Department of Science and Humanities

Random variables

- A Random variable is a set of **possible values** from a random experiment.

- A random variable X is a function from a sample space S into the real numbers R.

Types of random variables

- Random variables are classified into the following two categories:
- Discrete random variable
- Continuous random variable

Discrete random variables

- Discrete random variable: A random variable X is said to be discrete if its possible values form a discrete set. This means that if the possible values are arranged in order, then there is a gap between each value and the next one.
- The set of possible values may be infinite; for example, the set of all integers and the set of all positive integers are both discrete sets.
- For a discrete random variable, the possible values are the set of integers.

Probability Mass Function (PMF)

- For a discrete random variable X, if we specify the list of its possible values along with the probability that the random variable takes on each of these values, then we have completely described the population from which the random variable is sampled.
- For example, consider the following: The number of flaws in a 1-inch length of copper wire manufactured by a certain process varies from wire to wire. Overall, 48% of the wires produced have no flaws; 39% have one flaw; 12% have two flaws; and 1% have three flaws. Let X be the number of flaws in a randomly selected piece of wire.

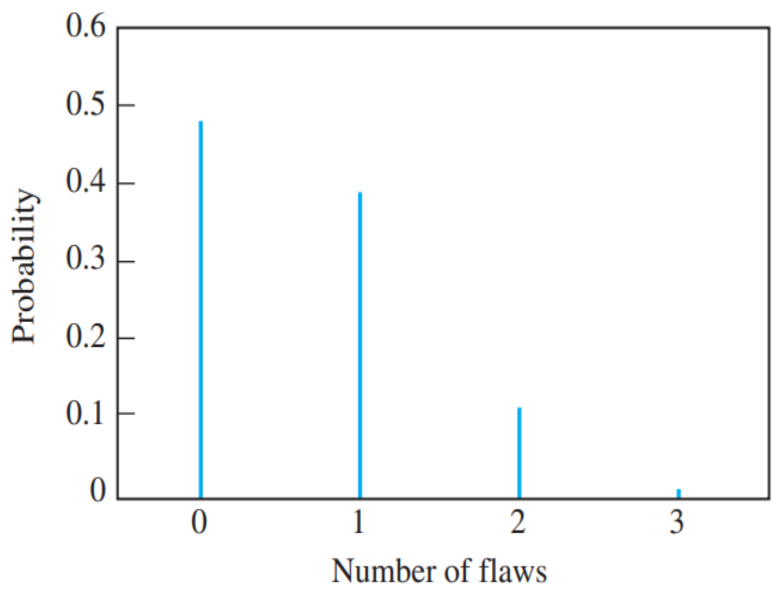
$$P(X = 0) = 0.48; P(X = 1) = 0.39; P(X = 2) = 0.12; P(X = 3) = 0.01$$

Probability Mass Function (PMF)

- The list of possible values $0,1,2,3,$ along with the probabilities for each, provide a complete description of the population from which X is drawn. This description has a name – the probability mass function.
- The probability mass function of a discrete random variable X is the function $p(x) = P(X = x)$. The probability mass function is sometimes called probability distribution.
- Thus for the random variable X representing the number of flaws in a length of wire, $p(0) = 0.48; p(1) = 0.39; p(2) = 0.12; p(3) = 0.01;$ and $\mathrm{p}(x) = 0$ for any value of x other than $0, 1, 2,$ or 3.

Probability Mass Function (PMF)

- Note that if the values of the probability mass function are added over all the possible values of X, the sum is equal to 1. This is true for any probability mass function.



This Figure presents a graph of the probability mass function of the random variable X . The physical interpretation of this graph is that each line represents a mass equal to its height.

Cumulative Distributive Function (CDF) of a Discrete Random Variable

- The probability mass function specifies the probability that a random variable is equal to a given value.
- A function called the cumulative distribution function specifies the probability that a random variable is less than or equal to a given value.
- The cumulative distribution function of the random variable X is the function $F(x) = P(X \leq x)$.
- Consider the following example. Let $F(x)$ denote the cumulative distribution function of the random variable X that represents the number of flaws in a randomly chosen wire. Find $F(2); F(1.5)$.

Cumulative Distributive Function (CDF) of a Discrete Random Variable

- By the definition of CDF, we have $F(x) = P(X \leq x)$.

  So, $F(2) = P(X \leq 2)$

  $= P(X = 0) + P(X = 1) + P(X = 2)$

  $= 0.48 + 0.39 + 0.12$

  $= 0.99$

Similarly, for $F(1.5)$,

$F(1.5) = P(X \leq 1.5) = P(X = 0) + P(X = 1) = 0.87$

Summary:

Let X be a discrete random variable. Then

- The probability mass function of X is the function $p(x) = P(X = x)$.
- The cumulative distribution of X is the function $F(x) = P(X \leq x)$.
- $\sum_x p(x) = \sum_x P(X = x) = 1$, where the sum is over all the possible values of X.

Mean and Variance for Discrete Random Variables

Let X be a discrete random variable with probability mass function $p(x) = P(X = x)$. Then.

- Mean of X is given by $\mu_X = \sum_x xP(X = x)$, where the sum is over all the possible values of X. The mean of X is sometimes called the expectation, or expected value, of X and may also be denoted by $E(X)$ or by μ.
- Variance of X is given by $\sigma_X^2 = \sum_x x^2 P(X = x) - \mu_X^2$
- Standard deviation is the square root of the variance: $\sigma_X = \sqrt{\sigma_X^2}$

Problem:

1. A random variable X has the following probability:

x	0	1	2	3	4	5	6
P(x)	k	3k	5k	7k	9k	11k	13k

Find i) $P(X < 4)$; ii) $P(X \geq 5)$; iii) $P(3 < X \leq 6)$.

Solution: If X is a random variable, then $\sum_{i=0}^{6} P(x_i) = 1 \Rightarrow k = \frac{1}{49}$

i) $P(X < 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 16k = \frac{16}{49}$

ii) $P(X \geq 5) = P(X = 5) + P(X = 6) = 24k = \frac{24}{49}$

iii) $P(3 < X \geq 6) = P(X = 4) + P(X = 5) + P(X = 6)$

$= 33k = \frac{33}{49}$

Department of Science and Humanities