Probability Theory

Probability Theory: Probability is defined as the chance of happening or occurrences of an event. Generally, the possibility of analyzing the occurrence of any event concerning previous data is called probability. For example, if a fair coin is tossed, what is the chance that it lands on the head? These types of questions are answered under probability. Probability measures the likelihood of an event's occurrence. In situations where the outcome of an event is uncertain, we discuss the probability of specific outcomes to understand their chances of happening. The study of events influenced by probability falls under the domain of statistics.

Basics of Probability Theory

Various terms used in probability theory are discussed below,

Random Experiment

In probability theory, any event which can be repeated multiple times and its outcome is not hampered by its repetition is called a Random Experiment. Tossing a coin, rolling dice, etc. are random experiments.

Sample Space

The set of all possible outcomes for any random experiment is called sample space. For example, throwing dice results in six outcomes, which are 1, 2, 3, 4, 5, and 6. Thus, its sample space is (1, 2, 3, 4, 5, 6)

Event

The outcome of any experiment is called an event. Various types of events used in probability theory are,

- Independent Events**:** The events whose outcomes are not affected by the outcomes of other future and/or past events are called independent events. **For example**, the output of tossing a coin in repetition is not affected by its previous outcome.

- Dependent Events**:** The events whose outcomes are affected by the outcome of other events are called dependent events. **For examp**le, picking oranges from a bag that contains 100 oranges without replacement.

- Mutually Exclusive Events**:** The events that can not occur simultaneously are called mutually exclusive events. **For example**, obtaining a head or a tail in tossing a coin, because both (head and tail) can not be obtained together.

- **Equally likely Events:** The events that have an equal chance or probability of happening are known as equally likely events. **For example**, observing any face in rolling dice has an equal probability of 1/6.

Random Variable

A variable that can assume the value of all possible outcomes of an experiment is called a random variable in Probability Theory. Random variables in probability theory are of two types which are discussed below,

Discrete Random Variable

Variables that can take countable values such as 0, 1, 2,… are called discrete random variables.

Continuous Random Variable

Variables that can take an infinite number of values in a given range are called continuous random variables.

Probability Theory Formulas

There are various formulas that are used in probability theory and some of them are discussed below,

- **Theoretical Probability Formula:** (Number of Favourable Outcomes) / (Number of Total Outcomes)
- **Empirical Probability Formula:** (Number of times event A happened) / (Total number of trials)
- **Addition Rule of Probability:** $P(A \cup B) = P(A) + P(B) – P(A \cap B)$
- **Complementary Rule of Probability:** $P(A') = 1 – P(A)$
- **Independent Events:** $P(A \cap B) = P(A) \cdot P(B)$
- **Conditional Probability:** $P(A \mid B) = P(A \cap B) / P(B)$
- **Bayes' Theorem:** $P(A \mid B) = P(B \mid A) \cdot P(A) / P(B)$

Discrete Random Variable:

Discrete Random Variables are an essential concept in probability theory and statistics. Discrete Random Variables play a crucial role in modelling real-world phenomena, from the number of customers who visit a store each day to the number of defective items in a production line. Understanding discrete random variables is essential for making informed decisions in various fields, such as finance, engineering, and healthcare.

In this article, we'll delve into the fundamentals of discrete random variables, including their definition, probability mass function, expected value, and variance. By the end of this article, you'll have a solid understanding of discrete random variables and how to use them to make better decisions.

Discrete Random Variable Definition

In probability theory, a discrete random variable is a type of random variable that can take on a finite or countable number of distinct values. These values are often represented by integers or whole numbers, other than this they can also be represented by other discrete values. For example, the number of heads obtained after flipping a coin three times is a discrete random variable. The possible values of this variable are 0, 1, 2, or 3.

Examples of a Discrete Random Variable

A very basic and fundamental example that comes to mind when talking about discrete random variables is the rolling of an unbiased standard die. An unbiased standard die is a die that has six faces and equal chances of any face coming on top. Considering we perform this experiment, it is pretty clear that there are only six outcomes for our experiment. Thus, our random variable can take any of the following discrete values from 1 to 6. Mathematically the collection of values that a random variable takes is denoted as a set. In this case, let the random variable be X.

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

Another popular example of a discrete random variable is the number of heads when tossing of two coins. In this case, the random variable X can take only one of the three choices i.e., 0, 1, and 2.

Other than these examples, there are various other examples of random discrete variables. Some of these are as follows:

- The number of cars that pass through a given intersection in an hour.
- The number of defective items in a shipment of goods.
- The number of people in a household.
- The number of accidents that occur at a given intersection in a week.
- The number of red balls drawn in a sample of 10 balls taken from a jar containing both red and blue balls.
- The number of goals scored in a soccer match.

Probability Distributions for Discrete Random Variables

The probability distribution of a discrete random variable is described by its probability mass function (PMF), which assigns a probability to each possible value of the variable. The key properties of a PMF are:

- Each probability is non-negative.
- The sum of all probabilities is equal to 1.

Common examples of discrete probability distributions include the binomial distribution, Poisson distribution, and geometric distribution.

Continuous Random Variable

Consider a generalized experiment rather than taking some particular experiment. Suppose that in your experiment, the outcome of this experiment can take values in some interval (a, b). That means that each and every single point in the interval can be taken up as the outcome values when you do the experiment. Hence, you do not have discrete values in this set of possible values but rather an interval.

Thus, $X= \{x: x \text{ belongs to } (a, b)\}$

Example of a Continuous Random Variable

Some examples of Continuous Random Variable are:

- The height of an adult male or female.
- The weight of an object.
- The time is taken to complete a task.
- The temperature of a room.
- The speed of a vehicle on a highway.

Random Variable

Random variable is a fundamental concept in statistics that bridges the gap between theoretical probability and real-world data. random variable in statistics is a function that assigns a real value to an outcome in the sample space of a random experiment. For example: if you roll a die, you can assign a number to each possible outcome.

There are two basic types of random variables,

- Discrete Random Variables (which take on specific values)
- Continuous Random Variables (assume any value within a given range)

What is a Random Variable?

Random Variable Probability is a mathematical concept that assigns numerical values to outcomes of a sample space. They can describe the outcomes of objective randomness (like tossing a coin) or subjective randomness(results of a cricket game).

There are two types of Random Variables- Discrete and Continuous.

A random variable is considered a discrete random variable when it takes specific, or distinct values within an interval. Conversely, if it takes a continuous range of values, then it is classified as a continuous random variable.

Random Variable Definition

Random variable in statistics is a variable whose possible values are numerical outcomes of a random phenomenon. It is a function that assigns a real number to each outcome in the sample space of a random experiment. We define a random variable as a function that maps from the sample space of an experiment to the real numbers. Mathematically, Random Variable is expressed as,

$X: S \to R$

where,

- X is Random Variable (It is usually denoted using capital letter)
- S is Sample Space
- R is Set of Real Numbers

Random Variable Example

Example 1

If two unbiased coins are tossed then find the random variable associated with that event.

Solution:

Suppose Two (unbiased) coins are tossed

X = number of heads. [X is a random variable or function]

Here, the sample space S = {HH, HT, TH, TT}

Example 2

Suppose a random variable X takes m different values i.e. sample space

X = {x1, x2, x3………xm} with probabilities

P(X = xi) = pi

where $1 \leq i \leq m$

The probabilities must satisfy the following conditions :

- $0 \leq pi \leq 1$; where $1 \leq i \leq m$
- p1 + p2 + p3 + ……. + pm = 1 Or we can say $0 \leq pi \leq 1$ and Σpi = 1

Hence possible values for random variable X are 0, 1, 2.

X = {0, 1, 2} where m = 3

- P(X = 0) = (Probability that number of heads is 0) = P(TT) = 1/2×1/2 = 1/4
- P(X = 1) = (Probability that number of heads is 1) = P(HT | TH) = 1/2×1/2 + 1/2×1/2 = 1/2
- P(X = 2) = (Probability that number of heads is 2) = P(HH) = 1/2×1/2 = 1/4

Here, you can observe that, ($0 \leq$ p1, p2, p3 $\leq 1/2$)

p1 + p2 + p3 = 1/4 + 2/4 + 1/4 = 1

For example,

Suppose a dice is thrown (X = outcome of the dice). Here, the sample space S = {1, 2, 3, 4, 5, 6}.

The output of the function will be:

- P(X=1) = 1/6
- P(X=2) = 1/6
- P(X=3) = 1/6
- P(X=4) = 1/6
- P(X=5) = 1/6

 P(X=6) = 1/6

Types of Random Variable

Random variables are of two types that are,

 Discrete Random Variable

 Continuous Random Variable

Discrete Random Variable

A Discrete Random Variable takes on a finite number of values. The probability function associated with it is said to be PMF.

PMF(Probability Mass Function)

If X is a discrete random variable and the PMF of X is P(xi), then

 $0 \leq pi \leq 1$

 Σpi = 1 where the sum is taken over all possible values of x

Discrete Random Variables Example

Example: Let S = {0, 1, 2}

x_i 0 1 2

$P_i(X = x_i)$ P_1 0.3 0.5

Find the value of P (X = 0)

Solution:

We know that the sum of all probabilities is equal to 1. And P (X= 0) be P1

P1 + 0.3 + 0.5 = 1

P1 = 0.2

Then, P (X = 0) is 0.2

Continuous Random Variable

Continuous Random Variable takes on an infinite number of values. The probability function associated with it is said to be PDF (Probability Density Function).

PDF (Probability Density Function)

If X is a continuous random variable. P (x < X < x + dx) = f(x)dx then,

 $0 \leq f(x) \leq 1$; for all x

 ∫ f(x) dx = 1 over all values of x

Then P (X) is said to be a PDF of the distribution.

Continuous Random Variables Example

Find the value of $P(1 < X < 2)$

Such that,

- $f(x) = kx3; 0 \le x \le 3 = 0$

Otherwise f(x) is a density function.

Solution:

If a function f is said to be a density function, then the sum of all probabilities is equal to 1.

Since it is a continuous random variable Integral value is 1 overall sample space s.

$\int f(x)\, dx = 1$

$\int kx^3\, dx = 1$

$K[x^4]/4 = 1$

Given interval, $0 \le x \le 3 = 0$

$K[3^4 – 0^4]/4 = 1$

$K(81/4) = 1$

$K = 4/81$

Thus,

$P(1 < X < 2) = k\times[X^4]/4$

$P = 4/81\times[16-1]/4$

$P = 15/81$

Random Variable Formulas

There are two main random variable formulas,

- Mean of Random Variable
- Variance of Random Variable

Let’s learn about the same in detail,

Mean of Random Variable

For any random variable X where P is its respective probability we define its mean as,

$\text{Mean}(\mu) = \Sigma X.P$

where,

- X is the random variable that consist of all possible values.
- P is the probability of respective variables

Variance of Random Variable

The variance of a random variable tells us how the random variable is spread about the mean value of the random variable. Variance of Random Variable is calculated using the formula,

$Var(x) = \sigma^2 = E(X^2) - \{E(X)\}^2$

where,

 $E(X^2) = \Sigma X^2P$

 $E(X) = \Sigma XP$

Random Variable Functions

For any random variable X if it assume the values $x_1, x_2, \ldots x_n$ where the probability corresponding to each random variable is $P(x_1), P(x_2), \ldots P(x_n)$, then the expected value of the variable is,

Expectation of X, $E(x) = \Sigma\, x.P(x)$

Now for any new random variable Y in which the random variable X is its input, i.e. Y = f(X), then the cumulative distribution function of Y is,

$F_y(Y) = P(g(X) \leq y)$

Random Variable Example with Solutions

Here are some of the solved examples on Random variable. Learn random variables by practicing these solved examples.

1. Find the mean value for the continuous random variable, $f(x) = x^2$, $1 \leq x \leq 3$

Solution:

Given,

$f(x) = x^2$

$1 \leq x \leq 3$

$E(x) = \int^3{}_1\, x.f(x)dx$

$E(x) = \int^3{}_1\, x.x^2.dx$

$E(x) = \int^3{}_1\, x^3.dx$

$E(x) = [x^4/4]^3{}_1$

$E(x) = 1/4\{3^4 - 1^4\} = 1/4\{81 - 1\}$

$E(x) = 1/4\{80\} = 20$

2. Find the mean value for the continuous random variable, $f(x) = e^x$, $1 \le x \le 3$

Solution:

Given,

$f(x) = e^x \; 1 \le x \le 3$

$E(x) = \int 3$

$1 \; x.f(x)dx$

$E(x) = \int^3_1 x.e^x.dx$

$E(x) = [x.e^x - e^x]^3_1$

$E(x) = [e^x(x - 1)]^3_1$

$E(x) = e^3(2) - e(0)$

3. Find the mean value for the continuous random variable, f(x) = x2, $1 \le x \le 3$

Solution :

f(x) = x^2, $1 \le x \le 3$

Mean = ∫(x^3)dx / ∫(x^2)dx from 1 to 3

$= [(1/4)x^4]_1^3 / [(1/3)x^3]_1^3$

= (81/4 – 1/4) / (27/3 – 1/3)

= 20 / 26/3

= 20 * 3/26 = 60/26 ≈ 2.31

4. Find the mean value for the continuous random variable, f(x) = 2x + 1, $0 \le x \le 4$

Solution :

f(x) = 2x + 1, $0 \le x \le 4$

Mean = ∫(x(2x+1))dx / ∫(2x+1)dx from 0 to 4

$= [(2/3)x^3 + (1/2)x^2]_0^4 / [x^2 + x]_0^4$

= (170/3) / 20

= 17/6 ≈ 2.83

5. Find the mean value for the continuous random variable, f(x) = x3, $-1 \le x \le 2$

Solution :

f(x) = x3, $-1 \le x \le 2$

Mean = ∫(x^4)dx / ∫(x^3)dx from -1 to 2

$= [(1/5)x^5]_{-1}^2 / [(1/4)x^4]_{-1}^2$

$= (32/5 + 1/5) / (16/4 + 1/4)$

$= 33/5 / 17/4$

$= 334 / (517) \approx 1.55$

6. Find the mean value for the continuous random variable, $f(x) = \sqrt{x}, 1 \leq x \leq 9$

Solution :

$f(x) = \sqrt{x}, 1 \leq x \leq 9$

Mean $= \int(x\sqrt{x})dx / \int(\sqrt{x})dx$ from 1 to 9

$= [(2/5)x^{(5/2)}]_1^9 / [(2/3)x^{(3/2)}]_1^9$

$= (486 - 2/5) / (18 - 2/3)$

≈ 5.2

7. Find the mean value for the continuous random variable, $f(x) = 3x2 - 2x, 0 \leq x \leq 3$

Solution :

$f(x) = 3x^2 - 2x, 0 \leq x \leq 3$

Mean $= \int(x(3x^2-2x))dx / \int(3x^2-2x)dx$ from 0 to 3

$= [(3/4)x^4 - (2/3)x^3]_0^3 / [x^3 - x^2]_0^3$

$= (81/4 - 18) / (27 - 9)$

$= 27/4 / 18 = = 27/72 = 3/8 \approx 0.375$

8. Find the mean value for the continuous random variable, $f(x) = \sin(x), 0 \leq x \leq \pi$

Solution :

$f(x) = \sin(x), 0 \leq x \leq \pi$

Mean $= \int(x \sin(x))dx / \int(\sin(x))dx$ from 0 to π

$= [-x \cos(x) + \sin(x)]_0^{pi} / [-\cos(x)]_0^{pi}$

$= (\pi + 1) / 2$

≈ 2.07

9. Find the mean value for the continuous random variable, $f(x) = ex, 0 \leq x \leq 2$

Solution :

$f(x) = e^x, 0 \leq x \leq 2$

Mean $= \int(x e^x)dx / \int(e^x)dx$ from 0 to 2

$= [x e^x - e^x]_0^2 / [e^x]_0^2$

$= (2e^2 - e^2 + 1) / (e^2 - 1)$

≈ 1.54

10. Find the mean value for the continuous random variable, $f(x) = \ln(x)$, $1 \leq x \leq e$

Solution :

$f(x) = \ln(x)$, $1 \leq x \leq e$

Mean = $\int(x \ln(x))dx / \int(\ln(x))dx$ from 1 to e

$= [(1/4)x^2(2\ln(x)-1)]_1^e / [x \ln(x) - x]_1^e$

$= (e^2/2 - 1/4) / (e - 1)$

≈ 1.95