## prepp

## Practice, Learn and Achieve Your Goal with Prepp

## NDA Exam

Study Material for Maths

## Simplifying <br> Government Exams

© SSC CHSL

© SSC CGL

(1) SBIPO
it ibps CLERK
妾AFCAT © SSCJE CTET
© CPIR UGC NET

## (1)CAPF

itz IBPS RRB

## PROBABILITY DISTRIBUTION

## Probability Distribution Definition

Probability distribution yields the possible outcomes for any random event. It is also defined based on the underlying sample space as a set of possible outcomes of any random experiment. These settings could be a set of real numbers or a set of vectors or set of any entities. It is a part of probability and statistics.

Random experiments are defined as the result of an experiment, whose outcome cannot be predicted. Suppose, if we toss a coin, we cannot predict, what outcome it will appear either it will come as Head or as Tail. The possible result of a random experiment is called an outcome. And the set of outcomes is called a sample point. With the help of these experiments or events, we can always create a probability pattern table in terms of variable and probabilities.

## Probability Distribution of Random Variables

A random variable has a probability distribution, which defines the probability of its unknown values. Random variables can be discrete (not constant) or continuous or both. That means it takes any of a designated finite or countable list of values, provided with a probability mass function feature of the random variable's probability distribution or can take any numerical value in an interval or set of intervals. Through a probability density function that is representative of the random variable's probability distribution or it can be a combination of both discrete and continuous.

Two random variables with equal probability distribution can yet vary with respect to their relationships with other random variables or whether they are independent of these. The recognition of a random variable, which means, the outcomes of randomly choosing values as per the variable's probability distribution function, are called random variates.

## Types of Probability Distribution

There are two types of probability distribution which are used for different purposes and various types of the data generation process.

- Normal or Cumulative Probability Distribution
- Binomial or Discrete Probability Distribution


## Cumulative Probability Distribution

The cumulative probability distribution is also known as a continuous probability distribution. In this distribution, the set of possible outcomes can take on values on a continuous range.

For example, a set of real numbers, is a continuous or normal distribution, as it gives all the possible outcomes of real numbers. Similarly, set of complex numbers, set of prime numbers, set of whole numbers etc. are the examples of Normal Probability distribution. Also, in real-life scenarios, the temperature of the day is an example of continuous probability. Based on these outcomes we can create a distribution table. A probability density function describes it. The formula for the normal distribution is;

$$
\mathrm{P}(\mathrm{x})=\frac{1}{\sigma \sqrt{2 \pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^{2}}
$$

Where,

- $\mu=$ Mean Value
- $\sigma=$ Standard Distribution of probability.
- If mean $(\mu)=0$ and standard deviation $(\sigma)=1$, then this distribution is known to be normal distribution.
- $x=$ Normal random variable


## Normal Distribution Examples

Since the normal distribution statistics estimates many natural events so well, it has evolved into a standard of recommendation for many probability queries. Some of the examples are:

- Height of the Population of the world
- Rolling a dice (once or multiple times)
- To judge Intelligent Quotient Level of children in this competitive world
- Tossing a coin
- Income distribution in countries economy among poor and rich
- The sizes of females shoes
- Weight of newly born babies range
- Average report of Students based on their performance


## Discrete Probability Distribution

A distribution is called a discrete probability distribution, where the set of outcomes are discrete in nature.

For example, if a dice is rolled, then all the possible outcomes are discrete and give a mass of outcomes. This is also known as probability mass functions.

So, the outcomes of binomial distribution consist of n repeated trials and the outcome may or may not occur. The formula for the binomial distribution is;

$$
\begin{aligned}
& \mathrm{P}(\mathrm{x})=\frac{n!}{r!(n-r)!} \cdot p^{r}(1-p)^{n-r} \\
& \mathrm{P}(\mathrm{x})=\mathrm{C}(\mathrm{n}, \mathrm{r}) \cdot p^{r}(1-p)^{n-r}
\end{aligned}
$$

Where,

- $n=$ Total number of events
- $r=$ Total number of successful events.
- $p=$ Success on a single trial probability.
- ${ }^{n} C_{r}=[n!/ r!(n-r)]!$
- 1 - $p=$ Failure Probability


## What is Negative Binomial Distribution?

In probability theory and statistics, if in a discrete probability distribution, the number of successes in a series of independent and identically disseminated Bernoulli trials before a particularised number of failures happens, then it is termed as the negative binomial distribution. Here the number of failures is denoted by ' $r$ '. For instance, if we throw a dice and determine the occurrence of 1 as a failure and all non-1's as successes. Now, if we throw a dice frequently until 1 appears the third time, i.e., $r=$ three failures, then the probability distribution of the number of non-1s that arrived would be the negative binomial distribution.

## What is Poisson Probability Distribution?

The Poisson probability distribution is a discrete probability distribution that represents the probability of a given number of events happening in a fixed time or space if these cases occur with a known steady rate and individually of the time since the last event. It was titled after French mathematician Siméon Denis Poisson. The Poisson distribution can also be practised for the number of events happening in other particularised intervals such as distance, area or volume. Some of the real-life examples are:

- A number of patients arriving at a clinic between 10 to 11 AM.
- The number of emails received by a manager between the office hours.
- The number of apples sold by a shopkeeper in the time period of 12 pm to 4 pm daily.


## Formulas

| $\mathrm{P}(\mathrm{X})={ }^{n} \mathrm{C}_{\mathrm{x}} \mathrm{a}^{\mathrm{x}} \mathrm{b}^{n-\mathrm{x}}$ <br> Where <br> Binomial Distribution $=$ probability of success <br> $\mathrm{b}=$ probability of failure <br> $\mathrm{n}=$ number of trials <br> $\mathrm{x}=$ random variable denoting success |  |
| :--- | :--- |
|  | $F_{X}(x)=\int_{-\infty}^{x} f_{X}(t) d t$ |
|  | $P(x)=\frac{n!}{r!(n-r)!} \cdot p^{r}(1-p)^{n-r}$ <br> $P(x)=C(n, r) \cdot p^{r}(1-p)^{n-r}$ |

## Probability Distribution Function

A function which is used to define the distribution of a probability is called a Probability distribution function. Depending upon the types, we can define these functions. Also, these functions are used in terms of probability density functions for any given random variable.

In the case of Normal distribution, the function of a real-valued random variable X is the function given by;
$\mathrm{F}_{\mathrm{X}}(\mathrm{x})=\mathrm{P}(\mathrm{X} \leq \mathrm{x})$
Where $P$ shows the probability that the random variable $X$ occurs on less than or equal to the value of $x$.


For a closed interval, $(\mathrm{a} \rightarrow \mathrm{b})$, the cumulative probability function can be defined as;
$P(a<X \leq b)=F_{X}(b)-F_{X}(a)$
If we express, the cumulative probability function as integral of its probability density function $f_{x}$, then,

$$
F_{x}(x)=\int_{-\infty}^{x} f_{x}(t) d t
$$

In the case of a random variable $X=b$, we can define cumulative probability function as;

$$
\mathrm{P}(\mathrm{X}=\mathrm{b})=\mathrm{F}_{\mathrm{x}}(\mathrm{~b})-\lim _{x \rightarrow b} f_{x}(t)
$$

## Probability Density Function Properties

Let $x$ be the continuous random variable with density function $f(x)$, the probability distribution function should satisfy the following conditions:

- For a continuous random variable that takes some value between certain limits, say a and b, and is calculated by finding the area under its curve and the $X$-axis, within the lower limit (a) and upper limit (b), then the pdf is given by $P(x)=\int_{a}^{b} f(x) d x$
- The probability density function is non-negative for all the possible values, i.e. $\mathbf{f}(\mathbf{x}) \geq \mathbf{0}$, for all x
- The area between the density curve and horizontal $X$-axis is equal to 1, i.e. $\int_{-\infty}^{\infty} f(x) d x=1$
- Due to the property of continuous random variable, the density function curve is continuous for all over the given range which defines itself over a range of continuous values or the domain of the variable.

Example: The pdf of a distribution is given as $f(x)\left\{\begin{array}{c}x ; \text { for } 0<x<1 \\ 2-x ; \text { for } 1<x<2 . \\ 0 ; \text { for } x>2\end{array}=\right.$
Calculate the density within the interval ( $0.5<x<1.5$ )

## Solution:

$$
\begin{aligned}
& \mathrm{P}(0.5<\mathrm{x}<1.5)=\int_{0.5}^{1.5} f(x) d x \\
& =\int_{0.5}^{1} f(x) d x+\int_{1}^{1.5} f(x) d x \\
& =\int_{0.5}^{1} x d x+\int_{1}^{1.5}(2-x) d x \\
& =\left(\frac{x^{2}}{2}\right)_{0.5}^{1}+\left(\left(2 x-\frac{x^{2}}{2}\right)\right)_{1}^{1.5} \\
& =3 / 4
\end{aligned}
$$

In the case of Binomial distribution, as we know it is defined as the probability of mass or discrete random variable gives exactly some value. This distribution is also called probability mass distribution and the function associated with it is called a probability mass function.

Probability mass function is basically defined for scalar or multivariate random variables whose domain is variant or discrete. Let us discuss its formula:

Suppose a random variable $X$ and sample space $S$ is defined as;

## $\mathbf{X}: \mathbf{S} \rightarrow \mathbf{A}$

And $A \in R$, where $R$ is a discrete random variable.
Then the probability mass function $f_{x}: A \rightarrow[0,1]$ for $X$ can be defined as;
$f_{\mathrm{X}}(\mathrm{x})=\mathrm{P}_{\mathrm{r}}(\mathrm{X}=\mathrm{x})=\mathrm{P}(\{\mathrm{s} \in \mathrm{S}: X(\mathrm{~s})=\mathrm{x}\})$

## Probability Distribution Table

The table could be created based on the random variable and possible outcomes. Say, a random variable $X$ is a real-valued function whose domain is the sample space of a random experiment. The probability distribution $\mathrm{P}(\mathrm{X})$ of a random variable X is the system of numbers.

| $X$ | $X_{1}$ | $X_{2}$ | $X_{3}$ | $\ldots \ldots \ldots \ldots \ldots$ | $X_{n}$ |
| :--- | :--- | :--- | :--- | :--- | :--- |
| $P(X)$ | $P_{1}$ | $P_{2}$ | $P_{3}$ | $\ldots \ldots \ldots \ldots \ldots$ | $P_{n}$ |

where $\mathrm{Pi}>0, \mathrm{i}=1$ to n and $\mathrm{P} 1+\mathrm{P} 2+\mathrm{P} 3+$ $\qquad$ $+\mathrm{Pn}=1$

## What is the Prior Probability?

In Bayesian statistical conclusion, a prior probability distribution, also known as the prior, of an unpredictable quantity is the probability distribution, expressing one's faiths about this quantity before any proof is taken into the record. For instance, the prior probability distribution represents the relative proportions of voters who will vote for some politician in a forthcoming election. The hidden quantity may be a parameter of the design or a possible variable rather than a perceptible variable.

## What is Posterior Probability?

The posterior probability is the likelihood an event will occur after all data or background information has been brought into account. It is nearly associated with a prior probability, where an event will occur before you take any new data or evidence into consideration. It is an adjustment of prior probability. We can calculate it by using the below formula:

## Posterior Probability = Prior Probability + New Evidence

It is commonly used in Bayesian hypothesis testing. For instance, old data propose that around 60\% of students who begin college will graduate within 4 years. This is the prior probability. Still, if we think the figure is much lower, so we start collecting new data. The data collected implies that the true figure is closer to $50 \%$, which is the posterior probability.

## Solved Examples

Example 1: A coin is tossed twice. $X$ is the random variable of the number of heads obtained. What is the probability distribution of $x$ ?

Solution: First write, the value of $X=0,1$ and 2 , as the possibility are there that
No head comes
One head and one tail comes
And head comes in both the coins
Now the probability distribution could be written as;
$P(X=0)=P($ Tail + Tail $)=1 / 2 * 1 / 2=1 / 4$
$P(X=1)=P($ Head + Tail ) or $P($ Tail+Head $)=1 / 2 * 1 / 2+1 / 2 * 1 / 2=1 / 2$
$P(X=2)=P($ Head + Head $)=1 / 2 * 1 / 2=1 / 4$
We can put these values in tabular form;

| $X$ | 0 | 1 | 2 |
| :--- | :--- | :--- | :--- |
| $P(X)$ | $1 / 4$ | $1 / 2$ | $1 / 4$ |

Example 2: The weight of a pot of water chosen is a continuous random variable. The following table gives the weight in $\mathbf{k g}$ of 100 containers recently filled by the water purifier. It records the observed values of the continuous random variable and their corresponding frequencies. Find the probability or chances for each weight category.

| Weight W | Number of Containers |
| :---: | :---: |
| $0.900-0.925$ | 1 |
| $0.925-0.950$ | 7 |
| $0.950-0.975$ | 25 |
| $0.975-1.000$ | 32 |
| $1.000-1.025$ | 30 |
| $1.025-1.050$ | 5 |
| Total | 100 |

## Solution:

We first divide the number of containers in each weight category by 100 to give the probabilities.

| Weight W | Number of Containers | Probability |
| :---: | :---: | :---: |
| $0.900-0.925$ | 1 | 0.01 |
| $0.925-0.950$ | 7 | 0.07 |
| $0.950-0.975$ | 25 | 0.25 |
| $0.975-1.000$ | 32 | 0.32 |
| $1.000-1.025$ | 30 | 0.30 |
| $1.025-1.050$ | 5 | 0.05 |
| Total | 100 | 1.00 |

## prepp

# Latest Sarkari jobs, <br> Govt Exam alerts, <br> Results and Vacancies 

> Latest News and Notification

- Exam Paper Analysis
- Topic-wise weightage
- Previous Year Papers with Answer Key
- Preparation Strategy \& Subject-wise Books

To know more Click Here

