SNS COLLEGE OF ENGINEERING

Kurumbapalayam(Po), Coimbatore – 641 107
Accredited by NAAC-UGC with 'A' Grade
Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

Department of Information Technology

19IT601– Data Science and Analytics

III Year / VI Semester

Unit 2 – DESCRIPTIVE ANALYTICS USING STATISTICS

Topic 2: Types of Data Distribution

Types of Data Distribution

1. Uniform Distribution
2. Normal Distribution
3. Exponential Distribution
4. Binomial Distribution
5. Poisson Distribution

Uniform Distribution

- A uniform distribution just means there's a flat constant probability of a value occurring within a given range.

- When you roll a fair die, the outcomes are 1 to 6. The probabilities of getting these outcomes are equally likely and that is the basis of a uniform distribution.

- all the n number of possible outcomes of a uniform distribution are equally likely.

- Every value, every range of values has an equal chance of appearing as any other value

Types of Data Distribution

Normal Distribution

It is otherwise known as Gaussian Distribution and Symmetric Distribution. It is a type of continuous probability distribution which is symmetric to the mean. The majority of the observations cluster around the central peak point.

Normal distribution represents the behavior of most of the situations in the universe.
The large sum of (small) random variables often turns out to be normally distributed, contributing to its widespread application.

Any distribution is known as Normal distribution if it has the following characteristics:

- The mean, median and mode of the distribution coincide.
- The curve of the distribution is bell-shaped and symmetrical about the line $x=\mu$.
- The total area under the curve is 1.
- Exactly half of the values are to the left of the center and the other half to the right.

Types of Data Distribution

Exponential Distribution

The exponential distribution is the probability distribution of the time between events in a Poisson point process, i.e., a process in which events occur continuously and independently at a constant average rate.

It is concerned with the amount of time until some specific event occurs.

Example:

- The amount of time until an earthquake occurs has an exponential distribution
- The amount of time in business telephone calls
- The car battery lasts.
- The exponential distribution is widely used in the field of reliability.

Types of Data Distribution

Binomial Distribution

- A binomial distribution can be thought of as simply the probability of a SUCCESS or FAILURE outcome in an experiment or survey that is repeated multiple times.
- The binomial distribution is used when there are exactly two mutually exclusive outcomes of a trial. These outcomes are appropriately labeled "success" and "failure“
- The binomial is a type of distribution that has two possible outcomes (the prefix “bi” means two, or twice).
- For example, a coin toss has only two possible outcomes: heads or tails and taking a test could have two possible outcomes: pass or fail.
- The terms p and q remain constant throughout the experiment, where p is the probability of getting a success on any one trial and q = (1 – p) is the probability of getting a failure on any one trial.

Types of Data Distribution

Poisson Distribution

- It is the discrete probability distribution of the number of times an event is likely to occur within a specified period of time. It is used for independent events which occur at a constant rate within a given interval of time.

- The occurrences in each interval can range from zero to infinity (0 to α).

Examples:

- How many black colours are there in a random sample of 50 cars
- No of cars arriving at a car wash during a 20 minute time interval

THANK YOU