# National Technology of Mexico
# Technological Institute of Tijuana

ACADEMIC SUBDIRECTION
Systems and Computing Department

SEMESTER
February – June 2021

CAREER
Information and Communication Technologies Engineer

SUBJECT AND KEY:
Data Mining    BDD-1703TI9A

STUDENT'S NAME AND REGISTRATION:
Velázquez Farrera César Alejandro    17212937

NAME OF THE JOB:
Statistical Distributions

UNIT TO BE EVALUATED
Unit II

TEACHER'S NAME:
Mc. José Christian Romero Hernández

## What is a probability distribution?

A probability distribution is a statistical function that describes all the possible values and likelihoods that a random variable can take within a given range. This range will be bounded between the minimum and maximum possible values, but precisely where the possible value is likely to be plotted on the probability distribution depends on a number of factors. These factors include the distribution's mean (average), standard deviation, skewness, and kurtosis.

## Types of probability distributions

There are many different classifications of probability distributions. Some of them include the normal distribution, chi square distribution, binomial distribution, and Poisson distribution. The different probability distributions serve different purposes and represent different data generation processes. The binomial distribution, for example, evaluates the probability of an event occurring several times over a given number of trials and given the event's probability in each trial. and may be generated by keeping track of how many free throws a basketball player makes in a game, where 1 = a basket and 0 = a miss.

The most commonly used distribution is the normal distribution, which is used frequently in finance, investing, science, and engineering. The normal distribution is fully characterized by its mean and standard deviation, meaning the distribution is not skewed and does exhibit kurtosis. This makes the distribution symmetric and it is depicted as a bell-shaped curve when plotted. A normal distribution is defined by a mean (average) of zero and a standard deviation of 1.0, with a skew of zero and kurtosis = 3. In a normal distribution, approximately 68% of the data collected will fall within +/- one standard deviation of the mean; approximately 95% within +/- two standard deviations; and 99.7% within three standard deviations. Unlike the binomial distribution, the normal distribution is continuous, meaning that all possible values are represented (as opposed to just 0 and 1 with nothing in between).

## Discrete Probability Distributions

Discrete probability functions are the probability of mass functions. It assumes a discrete number of values.

For example, when you toss a coin, then counts of events are discrete functions because there are no in-between values. You can have only heads or tails in a coin toss. Similarly, when counting the number of books borrowed per hour at a library, you can count 31 o 32 books and nothing in between.
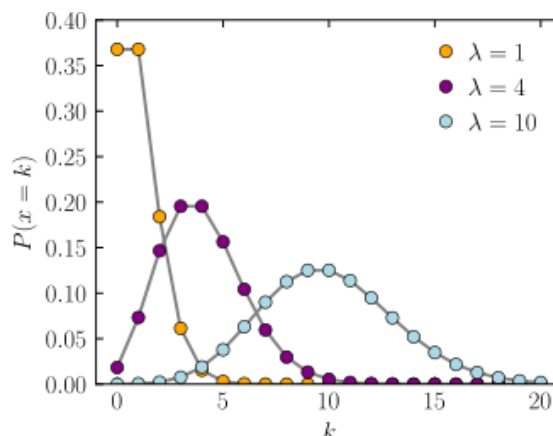
## Types of Discrete Probability Distributions

- **Binomial distributions**: A Bernoulli distribution has only two outcomes, 1 and 0. Therefore, the random variable X takes the value 1 with the probability of success as p, and the value 0 with the probability of failure as q or 1-p.
  Thus, if you toss a coin, the occurrence of head denotes success, and a tail denotes failure.
  The probability function is px(1-p)1-x where x € (0, 1)
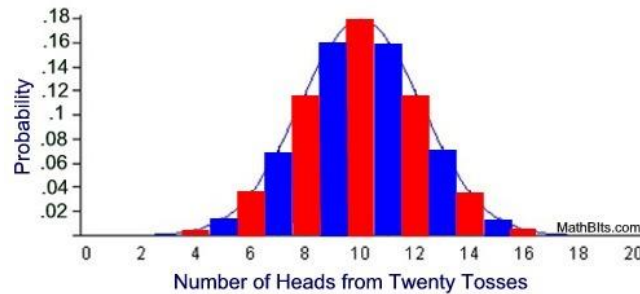


- **Normal distributions**: Normal distributions are for the most basic situations. It has the following characteristics:
  1. Mean, median and mode coincides.
  2. The distribution is bell-shaped
  3. The distribution curve is symmetrical along x = μ.
  4. The area under the curve is 1.
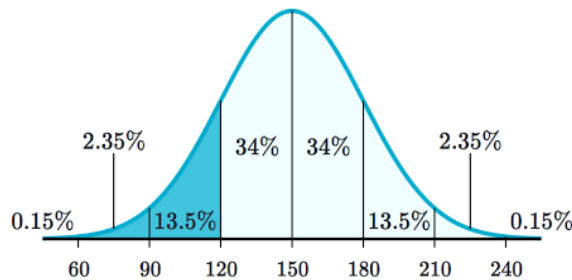


- **Poisson Distribution:** Counting number of books at a library falls under probability distribution. Poisson Distribution have the following assumptions:
  - A successful event is not influencing the outcome of another successful event.
  - The probability of success over a short duration equals the probability of success over a longer duration.
  - The probability of success in a duration nears zero as the duration becomes smaller.
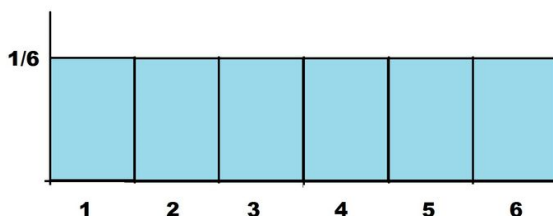
## Types of Continuous Probability Distributions

- **Uniform distributions** – When rolling a dice, the outcomes are 1 to 6. The probabilities of these outcomes are equal, and that is a uniform distribution.
  Suppose the random variable X assumes k different values. Also, P(X=xk) is constant.
  The P(X=xk) = 1/k



- **Cumulative probability distributions** - When the probability that the value of a random variable X is within a specified range, cumulative probability comes into the picture.
  Suppose you toss a coin, then what is the probability of the outcome to be one or fewer heads. This is a cumulative probability.

| Number of heads: x | Probability P(X=x) | Cumulative Probability: P(X <= x) |
|---|---|---|
| 0 | 0.25 | 0.25 |
| 1 | 0.50 | 0.75 |
| 2 | 0.25 | 1.00 |

References:

1. Pavan V. (2020). Types of Probability Distribution [Explained with Examples]. 4/29/2021. upGrad blog. Website: https://www.upgrad.com/blog/types-probability-distribution/
2. Adam H. (2020). Probability Distribution. 4/29/2021. Investopedia. Website: https://www.investopedia.com/terms/p/probabilitydistribution.asp

**Fecha de entrega:** 4/30/2021