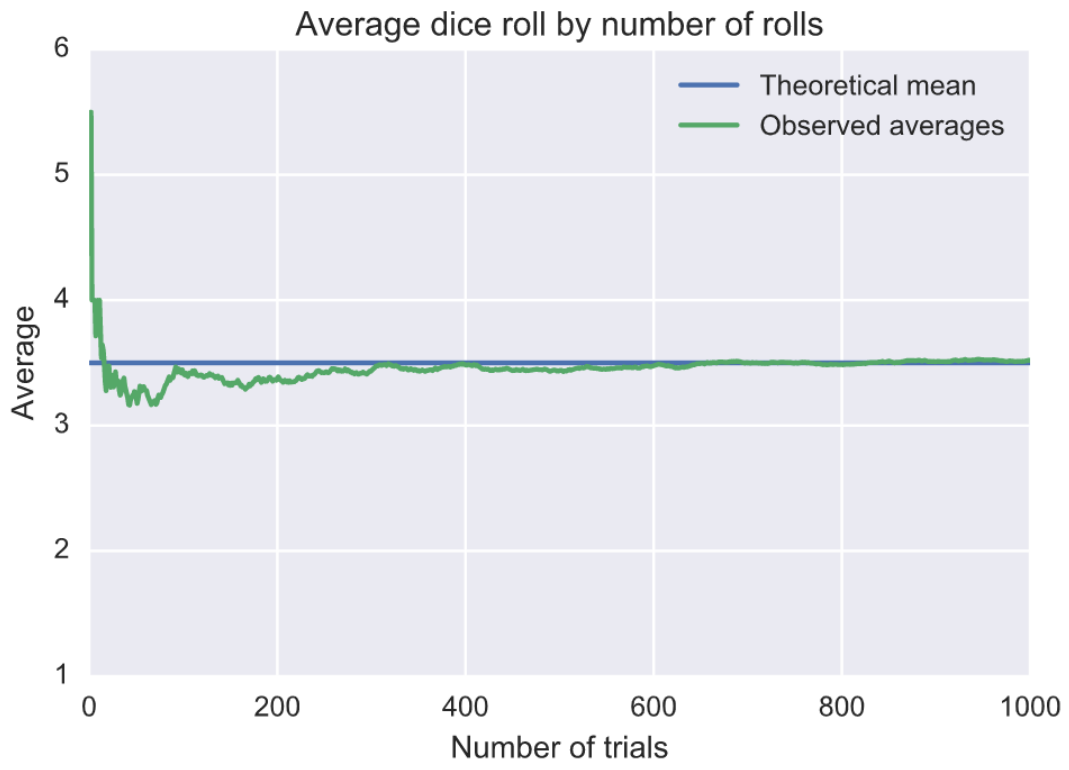
Probability Distributions

- Priyansu Panda

Law of Large Numbers



- In probability theory, the law of large numbers (LLN) is a theorem that describes the result of performing the same experiment a large number of times. According to the law, the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed.

Random Variables

- A random variable is a real-valued function whose domain is a set
- of possible outcomes of a random experiment and range is a sub-set of the set
- of real numbers and has the following properties:
- i) Each particular value of the random variable can be assigned some
- probability
- ii) Uniting all the probabilities associated with all the different values of the
- random variable gives the value 1(unity).

Discrete Random Variable

- A random variable is said to be discrete if it has either a finite or a countable number of values.

- The number of students present each day in a class during an academic session is an example of discrete random variable as the number cannot take a fractional value.

Probability Mass Function

Let X be a r.v. which takes the values x_1, x_2, ... and let $P[X = x_i] = p(x_i)$. This function $p(x_i)$, $i = 1, 2, \ldots$ defined for the values x_1, x_2, … assumed by X is called probability mass function of X satisfying $p(x_i) \geq 0$ and $\sum p(x_i) = 1$.

X	x_1	x_2	$x_3 \ldots$
$p(x)$	$p(x_1)$	$p(x_2)$	$p(x_3)\ldots$

X	0	1	2
$p(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Bernoulli Distribution

- There are experiments where the outcomes can be divided into two categories with reference to presence or absence of a particular attribute or characteristic.
- For example, head coming up in the toss of a fair coin may be treated as a success and tail as failure, or vice-versa.
- Accordingly, probabilities can be assigned to the success and failure.

The Bernoulli probability distribution, in tabular form, is given as

X	0	1
$p(x)$	$1-p$	p

Binomial Distribution

- binomial distribution which was discovered by J. Bernoulli (1654-1705) and was first published eight years after his death i.e. in 1713 and is also known as “Bernoulli distribution for n trials”.
- Binomial distribution is applicable for a random experiment comprising a finite number (n) of independent Bernoulli trials having the constant probability of success for each trial.

Binomial Distribution Definition

- A discrete random variable X is said to follow binomial distribution with parameters n and p if it assumes only a finite number of non-negative integer values and its probability mass function is given by

$$P[X = x] = \begin{cases} {}^{n}C_{x} p^{x} q^{n-x}; & x = 0, 1, 2, \ldots, n \\ 0; & \text{elsewhere} \end{cases}$$

- where, n is the number of independent trials,
- x is the number of successes in n trials,
- p is the probability of success in each trial, and
- q = 1 – p is the probability of failure in each trial.

Problem

A policeman fires 6 bullets on a dacoit. The probability that the dacoit will be killed by a bullet is 0.3. What is the probability that the dacoit is still alive?

Poisson Distribution

- Poisson distribution is a limiting case of binomial distribution under the following conditions:
- i) n, the number of trials is indefinitely large, i.e. n-> Infinity .
- ii) p, the constant probability of success for each trial is very small, i.e. p -> 0.
- iii) np is a finite quantity say ‘Lambda’.

Poisson Mass Function

- A random variable X is said to follow Poisson distribution if it assumes indefinite number of non-negative integer values and its probability mass function is given by:

$$p(x) = P(X = x) = \begin{cases} \dfrac{e^{-\lambda} . \lambda^{x}}{\lfloor x}; & x = 0, 1, 2, 3, \ldots \text{ and } \lambda > 0. \\ 0; & \text{elsewhere} \end{cases}$$

- where e = base of natural logarithm

Problem

If the probability that an individual suffers a bad reaction from an injection of a given serum is 0.001, determine the probability that out of 500 individuals

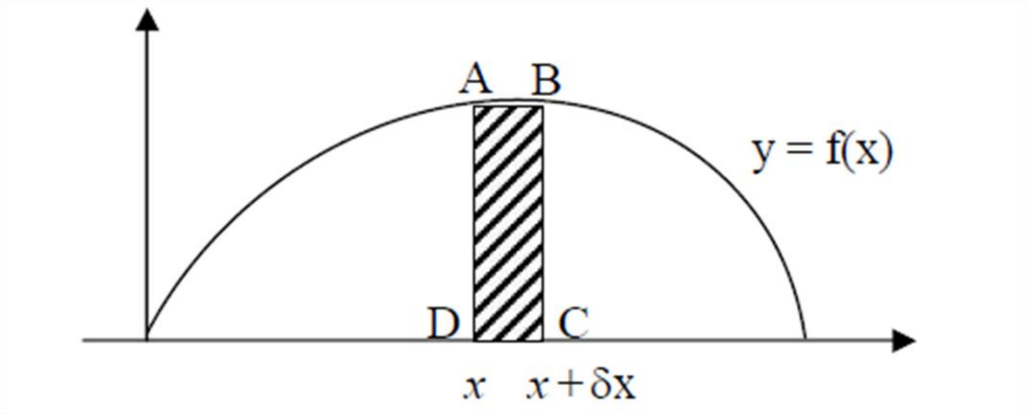
- i) exactly 3,
- ii) more than 2 individuals suffer from bad reaction

Continuous Random Variable

- Random variable is said to be continuous if it can take all possible real (i.e. integer as well as fractional) values between two certain limits.
- For example: Temperature of a city
- Rain fall difference between two cities

Probability Density Function

- Continuous random variable is represented by different representation known as **probability density function** unlike the discrete random variable which is represented by probability mass function



$$\int_{R} f(x)\,dx = 1,$$

Problem

- A continuous random variable X has the probability density function:

$$f(x) = Ax^3,\ 0 \le x \le 1.$$

Determine

i) A

ii) $P[0.2 < X < 0.5]$

Solution

(i) As f(x) is probability density function,

$$\therefore \int_R f(x)\,dx = 1$$

$$\Rightarrow \int_0^1 f(x)\,dx = 1 \Rightarrow \int_0^1 Ax^3 dx = 1$$

$$\Rightarrow A\left[\frac{x^4}{4}\right]_0^1 = 1 \Rightarrow A\left(\frac{1}{4} - 0\right) = 1 \Rightarrow A = 4$$

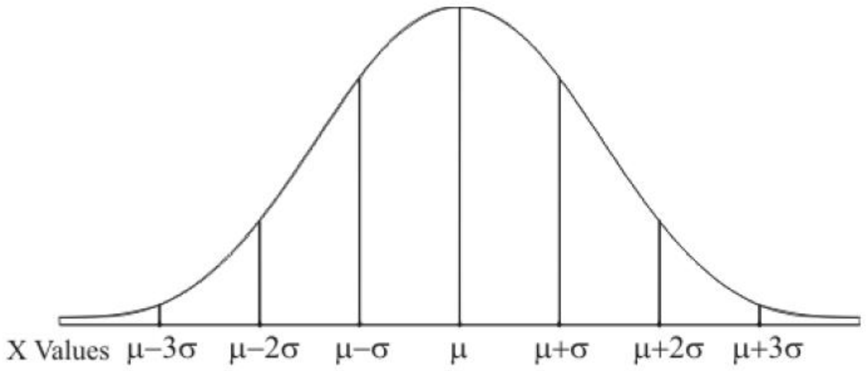
(ii) $$P[0.2 < X < 0.5] = \int_{0.2}^{0.5} f(x)\,dx = \int_{0.2}^{0.5} Ax^3 dx = 4\left[\frac{x^4}{4}\right]_{0.2}^{0.5} = [(0.5)^4 - (0.2)^4]$$

$$= 0.0625 - 0.0016 = 0.0609$$

Normal/Gaussian Distribution

- A continuous random variable X is said to follow normal distribution with parameters mean(mu) and standard deviation(sigma)if it takes on any real value and its probability density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \ -\infty < x < \infty;$$



- bell-shaped curve
- Mean, median and mode are same
- Symmetrical shape

Problem

In a university the mean weight of 1000 male students is 60 kg and standard deviation is 16 kg.

(a) Find the number of male students having their weights

i) less than 55 kg

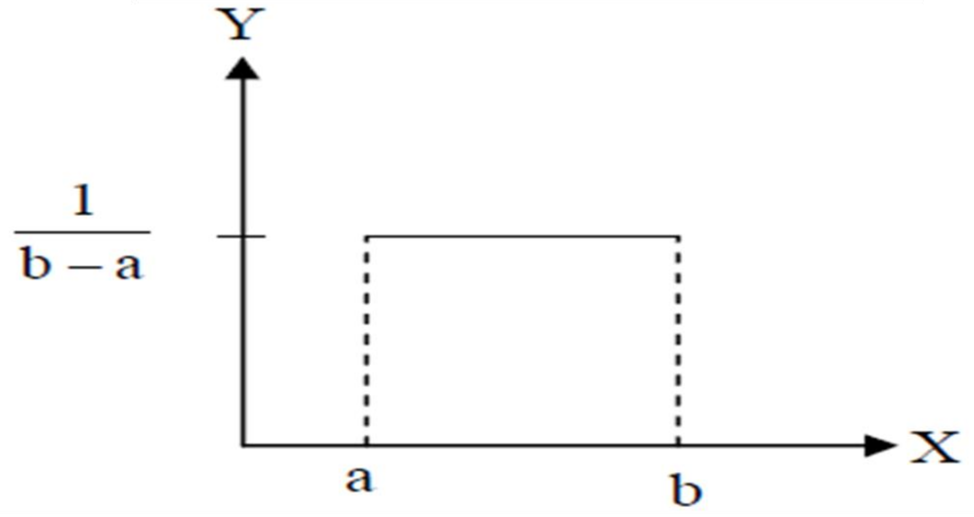
ii) more than 70 kg

iii) between 45 kg and 65 kg

Uniform Distribution

- A random variable X is said to follow a continuous uniform (rectangular) distribution over an interval (a, b) if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } a < x < b \\ 0, & \text{otherwise} \end{cases}$$

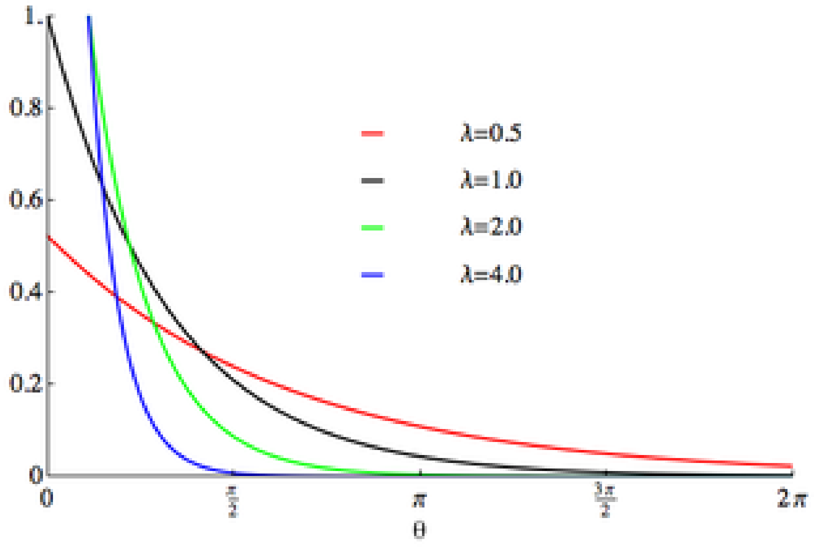


Exponential Distribution

- A random variable X is said to follow exponential distribution with parameter Lambda> 0, if it takes any non-negative real value and its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0 \\ 0, & \text{elsewhere} \end{cases}$$

- Lambda = 1/Mean
- Useful in modeling time difference between two events.



Problem

Suppose that accidents occur in a factory at a rate of 1/20 per working day. Suppose in the factory six days (from Monday to Saturday) are working and we begin observing the occurrence of accidents at the starting of work on Monday. Let X be the number of days until the first accident occurs. Find the probability that first week is accident free.